Journal of Experimental Social Psychology 112 (2024) 104588

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Moral thin-slicing: Forming moral impressions from a brief glance☆

Julian De Freitas [a,*,1], Alon Hafri [b,1]

[a] Harvard Business School, United States of America
[b] Department of Linguistics and Cognitive Science, University of Delaware, United States of America

ARTICLE INFO

Keywords:
Moral judgment
Thin slices
Event structure
Thematic roles
Visual relations

ABSTRACT

Despite the modern rarity with which people are visual witness to moral transgressions involving physical harm, such transgressions are more accessible than ever thanks to their availability on social media and in the news. On one hand, the literature suggests that people form fast moral impressions once they already know what has transpired (i.e., who did what to whom, and whether there was harm involved). On the other hand, almost all research on the psychological bases for moral judgment has used verbal vignettes, leaving open the question of how people form moral impressions about observed visual events. Using a naturalistic but well-controlled image set depicting social interactions, we find that observers are capable of 'moral thin-slicing': they reliably identify moral transgressions from visual scenes presented in the blink of an eye (< 100 ms), in ways that are surprisingly consistent with judgments made under no viewing-time constraints. Across four studies, we show that this remarkable ability arises because observers independently and rapidly extract the 'atoms' of moral judgment (i.e., event roles, and the level of harm involved). Our work supports recent proposals that many moral judgments are fast and intuitive and opens up exciting new avenues for understanding how people form moral judgments from visual observation.

1. Introduction

The ability to evaluate moral scenarios is a crucial human capacity that has significance both for everyday social interaction and for societal functioning at large. Thus, it is not surprising that the psychological bases for moral judgment have been extensively investigated. This includes research aiming to understand the informational, contextual, and cultural factors that serve as input to such judgments (De Freitas & Cikara, 2018; Greene et al., 2001; Haidt, 2012; Haidt et al., 1993; Strohminger & Victor, 2018), what kind of heuristics or biases are involved (De Freitas & Johnson, 2018; Gu et al., 2013; Haidt et al., 1993; Patil et al., 2017; Petrinovich & O'Neill, 1996; Wheatley & Haidt, 2005; Young & Saxe, 2009), and how much time and deliberation are necessary to come to a moral judgment (Cameron et al., 2017; Cusimano et al., 2017). While this research almost universally employs short verbal vignettes to investigate these questions (e.g., the trolley dilemma; Foot, 1967), in the current study, we seek to understand how people form moral impressions about observed visual events.

Given how much of our everyday social experience is rooted in our interaction with the visual world, it may at first be surprising that research on the mental processes involved in moral evaluation has generally used verbal vignettes rather than visual scenes (with a few exceptions: Decety & Cacioppo, 2012, Iliev et al., 2012, Nagel & Waldman, 2012, Caruso et al., 2016). However, upon reflection, this may be understandable. On the one hand, it is thankfully quite rare these days for individuals to directly witness moral transgressions — at least those involving physical harm — compared with past eras (Pinker, 2012). Second, and more practically speaking, verbal vignettes are an optimal stimulus for isolating the factors that might contribute to moral judgments. Indeed, one inherent property of language is that it furnishes a framing of an event that is often relatively unambiguous: a speaker's perspective on an event will determine how they describe it (e.g., as "kill" or "cause to die"; De Freitas et al., 2017), and this will in turn shape the listener's inferences about it (Jackendoff, 2010). Thus, vignettes efficiently and explicitly 'package' and communicate information such as the degree of causation, harm, and intentionality attributable to event participants (or lack thereof).

On the other hand, video and social media have made moral

☆ This paper has been recommended for acceptance by Dr. Paul Conway
* Corresponding author.
E-mail address: jdefreitas@hbs.edu (J. De Freitas).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jesp.2023.104588
Received 24 August 2023; Received in revised form 29 December 2023; Accepted 30 December 2023

transgressions more accessible visually than ever before.[2] Such transgressions are of course depicted in television and film, but in fact, visual examples of real-world transgressions are ubiquitous (e.g., the examples in Fig. 1a): they can be captured, uploaded online, and distributed via social networks extending from the present into the future; aggregated in memes on social media and video-sharing sites (as in hashtags like #instant karma, #random acts of kindness, #dashcam scam, and others); and covered in the latest news online. A single video can even start a social movement, as when the viral video of a police officer killing a Black US citizen named George Floyd catalyzed global protests.[3]

While the importance of evaluating the moral content of observed events is clear, it is currently unknown how morally relevant information is extracted in the case of visually observed transgressions of one actor engaging with another. Of interest in the current study is how much exposure to a visual event — without the support of descriptive verbal information (e.g., headlines or other commentary commonly found in news and online content) — is sufficient for observers to form moral judgments that are consistent from one observer to the next. In particular, the focus of the current study is whether observers can form such impressions rapidly, after just a brief glance and without verbal context — a capability we call 'moral thin-slicing'.

1.1. *Thin-slicing in a moral context*

The ability to make judgments based on brief periods, or 'thin slices', of information is referred to as 'thin-slicing' in the context of psychology (Ambady & Rosenthal, 1992; Peracchio & Luna, 2006). Such work has explored the role that fast and automatic processes play in the decisions that viewers make (e.g., when talking with a customer service representative), including the effects of nonverbal cues such as facial expression and tone of voice.

Although people can make sub-second 'thin-slice' judgments in some contexts, there are reasons to think that they might not be able to do so for *moral* judgments, given how complex and multi-faceted such judgments are. In particular, moral judgments involve the evaluation of a variety of different properties, many of which are quite subjective. To make a moral judgment of the individuals in an interaction, at minimum one needs to know who in the interaction is the doer (aka 'Agent') or recipient (aka 'Patient') and whether the action is bad or harmful. To add to the complexity of such evaluations, there are also other factors that affect moral judgment, such as the intentional states of the participants (Allison et al., 2000; Cushman, 2008; Dennett, 1989), the causal factors leading up to the event (De Freitas et al., 2017; Tsiros et al., 2004; Weiner, 2000), and whether the agent had justifiable reasons for their actions (Malle et al., 2014). What is more, there is debate as to whether 'moral judgment' is even a single psychological notion at all; instead, it may involve distinct types of judgments, including basic evaluations of good vs. bad, judgments about social norms (e.g., forbidden vs. permissible), judgments of moral wrongness, and judgments of blame (see Malle, 2021, for a review and synthesis of the recent literature).

Nevertheless, despite the complex and subjective nature of moral evaluation processes, several lines of evidence point to the ability of people to form moral impressions rapidly from just a minimal amount of information. The 'social intuitionist' model, for instance, says that people's moral judgments are often informed by a fast, intuitive (rather than a deliberative and comprehensive) assessment of a situation (Haidt et al., 2000). For example, people's moral judgments of whether it is okay for two siblings to engage in sexual intercourse is driven more by their intuitive feelings of disgust than by reasoned thinking about whether anyone was harmed (Haidt et al., 1993). Other more recent proposals similarly advocate for fast and intuitive notions of 'purity' (Graham et al., 2013) or harm (Schein & Gray, 2018) in these assessments. Additionally, prior work has found that people can quickly categorize the harm and/or moral valence of different verbally presented actions, often in less than one second (e.g., "stealing", "gossip"; Schein & Gray, 2015, Cameron et al., 2017).

1.2. *The computational challenge of extracting morally relevant information from visual observation*

What all of the above studies have in common is their use of verbal stimuli to probe processes of moral evaluation, leaving open the question of how such computations occur over visual scenes. Crucially, it is not a given that fast and intuitive judgments based on linguistically presented scenarios would carry over to processes involved in interpreting visually observed events. Indeed, in most cases, participants in studies using verbal stimuli will have already read and understood what has transpired — i.e., they will have learned who acted on whom and what the actors did — before having to evaluate the scenario. By contrast, upon encountering a visual scene (like those in Fig. 1a), observers must build an interpretation of the event 'from scratch', identifying morally relevant aspects (such as causation, harm, intentionality, agency, and context) from the image itself.

Indeed, moral judgment poses a formidable computational challenge for visual perception because moral judgments require extracting not only the actions or visual properties of a single individual, but also the relations *between* individuals. Thus, moral judgments of visually depicted social interactions involve integrating inferences about individuals into a whole (Hafri & Firestone, 2021). Furthermore, there is wide variation in what a harmful action can look like in terms of both its fine-grained details and coarse-grained postural information.

Additionally, making moral evaluations of observed events is complex because it involves integrating information from multiple mental systems spanning vision (which is generally fast and automatic) and cognition (which is often relatively slow and deliberative) (Kahneman, 2011). The visual system is able to rapidly detect whether an interaction is social or non-social based on whether individuals are facing toward or away from one another (Isik et al., 2017; Papeo, 2020; Papeo & Abassi, 2019). By contrast, since there are no literal features in an image that trivially indicate moral wrongness, cognitive brain areas beyond the visual system, such as the ventromedial prefrontal cortex, might reasonably be required for integrating together abstract information inferred from a visual scene to make moral judgments (Greene & Haidt, 2002).

1.3. *Moral 'Atoms' and the possibility of moral thin-slicing*

Here we investigate whether observers can make a moral judgment about a briefly viewed visual scene by rapidly extracting and assembling together the 'atoms' of moral judgment: event role information (who acted on whom) and harm level (whether an action caused harm or did not).[4]

[2] Indeed, in an online survey, we have found that when it comes to the news or online media, moral transgressions involving physical harm (e.g., *slapping*) are witnessed almost as frequently as events involving little or no harm (e.g., *hugging*) — even as such transgressions are witnessed comparatively rarely in a direct, firsthand context. See the Supplementary Material for more details on this "pre-study."

[3] https://en.wikipedia.org/wiki/List_of_George_Floyd_protests_outside_the_United_States

[4] As a first pass look at whether moral information can be extracted from a brief glance, we operationalize a moral transgression as whether an act involved a low or high degree of physical harm; thus, whoever is the active participant in such an event — namely, the agent — would be the one considered to be morally wrong. However, we recognize that different kinds of moral judgments involve integrating many more kinds of information beyond just event roles and harm level, including intentionality, norm violations, and 'badness' (for nuanced perspectives on these issues, see Schein & Gray, 2018; Malle, 2021). We return to these issues in the General Discussion.

Fig. 1. (a) Individuals often view images of many kinds of low- and high-harm interactions every day in social media, news, video games, and other simulated environments. Examples of such images are depicted here. (b) Controlled image set used in our experiments, in which identical twins took part in low- and high-harm interactions, with actor side and event role fully counterbalanced. Sources for images in (a), from left-to-right by row: https://i.ytimg.com/vi/vahi77oOsK4/maxresdefault.jpg; https://touringplans.com/blog/face-5-things-didnt-know-face-painting-disney-world/; https://minnesota.cbslocal.com/wp-content/uploads/sites/15909630/2021/06/Screen-Shot-2021-06-11-at-1.34.56-PM.png; https://chicago.suntimes.com/2017/4/27/18358404/united-settles-with-dr-david-dao-who-was-dragged-from-jet.)

How rapid might this process be? One possibility is that people must view and deliberate over visual scenes for *at least several seconds* to judge a moral transgressor. This would show that such moral judgments rely solely on relatively lengthy, reflective processes, akin to predicting a negotiation outcome. Indeed, even many studies under the banner of 'thin-slicing' have actually asked people to evaluate visual content presented not for mere milliseconds but for several seconds or even minutes (Ambady & Rosenthal, 1992), such as predicting a negotiation outcome based on the first five minutes of the interaction (Curhan & Pentland, 2007), or evaluating the trustworthiness of a salesperson based on a 30-s clip (Ambady et al., 2006; Hall et al., 2015; Main et al., 2007).

A second, more exciting, possibility is that people need to observe a social interaction for *less than a blink of an eye* (< 100 ms), which would show that some moral judgments rely on relatively rapid, automatic processes, more akin to judging the trustworthiness of a face or distinguishing an ad from an editorial (Pieters & Wedel, 2012; Willis & Todorov, 2006). Providing some tentative evidence toward this possibility, previous work has found that observers rapidly detect the atoms on which moral judgment depends. One study found that people presented with naturalistic photographs of social interactions needed just 37 ms of viewing time to accurately recognize the roles of the actors involved (i.e., Agent or Patient) by processing relative configural shape features, e.g., whether one person is leaning forward and has his extremities outstretched more than the other person (Hafri et al., 2013). A more recent study showed that observers extract this information even when engaged in orthogonal, trivially simple tasks (e.g., color identification), suggesting a measure of automaticity to such extraction (Hafri et al., 2018). Another study found that visually processed information about irrelevant causal events affected peoples' moral judgments even when people believed they did not. De Freitas and Alvarez (2018) showed participants events in which it was ambiguous whether (a) one car hit into another car, which in turn hit into a pedestrian, or (b) the second car accelerated on its own into the pedestrian. By including simple, task-irrelevant peripheral events (i.e., two discs colliding) that participants did not believe affected their judgments, the main display was made to look more causal, which in turn increased moral blame attribution for the driver of the first car.

As for information about harm, it is currently unknown whether information about harm is extracted and consciously accessible from just a brief glance. However, previous work has found that people may rapidly extract the 'gist' of a social interaction, i.e., its basic-level category (e.g., 'kicking' or 'tickling'; Hafri et al., 2013), after brief exposure, suggesting that they should also be able to rapidly categorize events as harmful or not harmful from such exposures. The most relevant work to this issue is a pair of studies by Decety and colleagues that explored the locus and time course of brain responses evoked when observers evaluated visual stimuli that varied in degree and type of harm: moral and nonmoral transgressions (e.g., one person shoving another, a person accidentally knocking down a mug, etc.; (Decety & Cacioppo, 2012) or social scenes involving either harm or assistance (e.g., a person pulling another's hair, or helping someone off the floor, etc.; Yoder & Decety, 2014). The authors found that electrophysiological responses could be used to differentiate the intentionality and harm of an action within a few hundred milliseconds. Crucially, however, this work did not assess the conscious availability of such harm information at this latency, nor did it experimentally manipulate or counterbalance the roles of the participants (i.e., having each participant play the Agent and Patient role in the same event type across stimuli) — a necessary experimental design for revealing whether the mind encodes and utilizes this particular moral atom (event roles) in evaluating the actors involved.

1.4. Integrating moral atoms extracted perceptually

Even if information about event role and harm is available to an observer after a brief glance, it is currently unknown whether observers can rapidly *integrate* these disparate types of morally relevant event information to make a moral judgment. Indeed, it is entirely possible that information about these moral atoms, despite being extracted at rapid speeds, is cognitively segregated and is only integrated when observers reflect about an event deliberatively and effortfully. This situation could arise, for example, if the two types of information are output in different representational formats (e.g., imagistic vs. abstract; Marr, 1982), or are processed by two different visual streams (Milner & Goodale, 2006; Ungerleider & Mishkin, 1982).

Information segregation is regularly evidenced in both visual and cognitive processing, as in the "what"/"where" (or "what"/"how") division in visual processing (Milner & Goodale, 2006, Ungerleider & Mishkin, 1982). Indeed, patients with damage to the ventral stream show selective deficits in representing object information but have intact object localization or action information, while those with damage to the dorsal stream show the reverse effects (Milner & Goodale, 2006, Ungerleider & Mishkin, 1982). Even neurotypical individuals exhibit subtle signs of this pattern, as when they experience a perceptual size illusion while still accurately adjusting their fingers to the correct size of the misperceived object when reaching for it (Rossetti, 1998). Thus, information represented in a segregated manner in the mind can fail to be integrated successfully toward a common behavioral goal.

Another example of segregated information not being readily integrated comes from the literature on spatial navigation. In certain experimental tasks, participants are disoriented in unfamiliar environments and must reorient to find rewards. Crucially, the environments are designed such that geometric information (i.e., the overall 'shape' of the environment) is insufficient to disambiguate the correct location; nongeometric information is also required (e.g., the color or texture at certain locations). Young children and many nonhuman animals fail to use the nongeometric cues, instead relying primarily on local geometry to reorient: they search not only in the correct location but also in rotationally equivalent locations, even when those locations are readily distinguishable based on nongeometric cues that the animals are sensitive to (Cheng, 1986; Hermer & Spelke, 1994; Julian et al., 2015). Strikingly, adults under linguistic interference (i.e., verbal shadowing) show similar error patterns, despite being able to detect and remember both types of information (Hermer-Vazquez et al., 1999).

Collectively, these examples demonstrate situations in which disparate kinds of visual information fail to be integrated successfully toward a common behavioral goal, i.e., perceiving and grabbing an object, or spatially reorienting using geometric and non-geometric cues. Relating to the current investigation, while observers of social interactions may rapidly extract role and harm information, they may not automatically integrate them to make moral judgments — perhaps because the information remains cognitively segregated unless actively deliberated upon.

1.5. 'Moral thin-slicing' from visual observation?

In contrast to the possibility that distinct pieces of morally relevant information extracted from a scene remain cognitively segregated, here we hypothesize that observers are capable of 'moral thin-slicing' from visual observation. In other words, they can identify who the moral transgressor is in images presented for less than 100 milliseconds in a way consistent with those made for the same images viewed without time constraints. Furthermore, this process should be rapid, enabling them to report these judgments within a mere second or two at most.

We further propose that the success of such moral thin-slicing depends on a 'perceptual bottleneck': how quickly people can extract the inputs (the 'atoms') on which moral judgment depends, i.e., the different roles of the event participants and whether the interaction was harmful or not. By the same token, causally increasing how challenging it is to extract event role and harm information should lead to decrements in the ability of observers to make consistent moral judgments.

1.6. The present studies

To determine whether people are truly capable of moral thin-slicing without other confounding factors, such as different scene contexts, social identities (e.g., bouncer vs. patron), or viewpoints, our studies employed a controlled set of photographs (Fig. 1b). Given images of actors engaged in either low- or high-harm social interactions, participants were tasked with determining who acted on whom ('Role'), whether harm was inflicted ('Harm'), and whether each actor was doing something morally wrong ('Moral Wrongness').[5] Study 1 ('Moral Thin-Slicing') established the basic moral thin-slicing effect, Study 2 ('Temporal Evolution of a Moral Judgment') traced the evolution of a moral judgment from 17 ms to 1500 ms, and Study 3 ('Causal Manipulation of Role') and Study 4 ('Causal Manipulate of Harm') causally intervened on these psychological processes.

All experiments were approved by the Institutional Review Boards at Harvard Business School and the University of Delaware. Anonymized trial-level data for all experiments reported in this manuscript are publicly available on the Open Science Framework and are accessible at https://osf.io/qegjc/.

2. Study 1: Moral thin-slicing

Study 1 tested whether participants can detect moral information when presented in the blink of an eye. To do so, we first ran a separate 'unspeeded' norming study ($n = 169$, Study S1 in Supplementary Material) in which there were no viewing-time constraints, which allowed us to carefully norm the stimuli for the main experiments. The stimulus set was well suited for precisely probing extraction of properties that contribute to moral judgment (event role and harm level), as the spatial location and identity of the agent for each social interaction category were fully counterbalanced while other factors were kept constant (e.g., the same neutral scene context was always used).

2.1. Method

2.1.1. Participants

We collected data from 134 participants from Amazon's Mechanical Turk (Mturk), with the aim of achieving sample sizes in each task condition on par with previous studies using similar stimuli and tasks (i.e., approximately $n = 24$; Hafri et al., 2013, Hafri et al., 2018, Hafri et al., 2022). Sample sizes for this and all other studies were determined before any data analysis. Sample sizes and analysis plans were not preregistered.

The study link specified that participants should have normal or corrected-to-normal visual acuity. Participants in all studies were allowed to participate if they had an Mturk approval rating above 95%, had participated in at least 100 studies previously, were located in the USA, and had not taken part in any previous study in this project. These prescreening criteria were designed to select for motivated online study participants who were naïve to the purposes of the studies. We additionally included several attention checks and monitored the timing of stimuli to ensure reliability of data. We established additional exclusion criteria based on these attention checks and timing characteristics (outlined below in the section entitled *Exclusions*). Thirty-one participants were excluded from analysis based on these exclusion criteria, leaving data from 105 participants for analysis (57 identifying as male, 48 as female; mean age 34.2, *sd* 10.5, range 18–70).

We took no special measures beyond these for recruiting participants, and thus the sample was as diverse and inclusive as the population of online participants that met the above recruitment criteria at the time of data collection, which is generally more diverse than the typical college undergraduate (Mason & Suri, 2012; Peer et al., 2017). Beyond the above criteria, we have no reason to believe that the results reported here depend on characteristics of the participants not considered above.

2.1.2. Stimuli

Data from the norming study (Study S1) were used to curate a set of social interaction images for which participants provided consistent judgments about morally relevant information, given no time constraints on viewing time (see Supplementary Material for more details). In all studies reported in this paper, we employed 108 images of identical-twin actors engaged in 27 common social interaction categories varying in the degree of harm. Based on results from norming, fifteen of these interactions were classified as 'low harm' (bandaging, brushing, calling after, feeding, cover, poking, lifting, tickling, look at, face-painting, hugging, kissing, dressing, tapping, filming); nine as 'high harm' (strangling, shooting, kicking, punching, slapping, tripping, scratching, stabbing, biting); and three were excluded according to criteria described in the *Exclusions* section below.

The norming study allowed us to code various aspects of the images in a binary manner: for Color, which individual was wearing the red (or blue) shirt; for Role, which individual was the Agent or Patient; for Harm, whether harm was high or low; and for Moral Wrongness, whether each individual in the event was doing something morally wrong or not. We note that in our image set, in all cases where an action was considered harmful, the Agent was also considered to be doing something morally wrong. We used these binary categories for our analyses (although we recognize that some of these properties may nevertheless be represented psychologically in a graded fashion).

For each social interaction category, we counterbalanced the spatial location of the agent as well as the colors of their shirts (blue or orange-red, so that they would appear distinct even to color-blind individuals). Otherwise, the actors were similar in all respects: they were identical twins (age 29) with similar haircuts and clothes (apart from shirt color). The actors were photographed against a plain light-blue background, and then the background was post-processed to a uniform level of brightness. In images where an instrument was an important part of the action (e.g., a knife, for the stabbing category), both the actor (aka 'agent') and recipient (aka 'patient') held duplicates of the instrument, in order that possession of the instrument on its own would not allow the observer to determine event roles. For examples of these images, see Fig. 1b.

2.1.3. Procedure

After consenting, participants were redirected to a web server where platform-independent stimulus presentation and data collection were completed by custom software run in their web browsers, written using a combination of html, CSS, and jQuery.

Each participant saw all the images, but the task they completed in response to these images varied between-subjects, consisting of the following questions: (i) *Color*: "Was the person on the LEFT (RIGHT) wearing a red shirt?", (ii) *Role*: "Was the person on the LEFT (RIGHT) acting on the other person?", (iii) *Harm*: "Was there harm being inflicted?", (iii) *Moral Wrongness*: "Was the person on the LEFT (RIGHT) doing something morally wrong?" For the moral wrongness task, participants were asked about the agent for half of the images, and the patient for the other half. For all tasks, the answer options were always "yes" or "no". Which side (left or right) participants were asked about was counterbalanced within-subjects for the color, role, and moral wrongness tasks. The color task was intended as a non-social baseline for which we

[5] Since this is the first study (to our knowledge) to seek behavioral evidence for moral thin-slicing from visual observation in a systematic manner, we narrowed our focus to intentional actions with or without physical harm; thus, the main properties participants had to extract to make a moral judgment were (1) whether the action was harmful or not, and (2) the identity of the agent and patient. However, future work could explore the impact of intentionality or norm violations on such judgments — a possibility to which we return in the General Discussion (for nuanced perspectives on these issues, see Schein & Gray, 2018; Malle, 2021).

expected participants to perform well. For the moral wrongness task, we expected agents and patients to elicit categorically distinct moral judgments (Gray et al., 2012), so for this task we aimed for twice as many participants as in the other conditions (i.e., approximately $n = 48$) in order to maintain similar statistical power for judgments about each event role.

Because the display loaded within participants' own web browsers, viewing distance and screen resolutions could vary dramatically, so we report dimensions of the stimuli using pixel (px) values and positions of the stimuli as pixel values relative to the left and top borders of a gray (red [R]: 221, green [G]: 221, blue [B]: 221) task window (800 px × 554 px), within which the images were presented.

The gray task window was always present on the screen. At the beginning of a trial, the word "Ready?" (font size: 14 pt) appeared for 400 ms in the middle of the task window, followed by a fixation cross (font size: 14 pt) that appeared for 100 ms in the same location, followed by the image (590px × 443px; left: 105px, top: 3px) which was shown for 33 ms before disappearing. The task question (615 px × 10 px; left: 100 px, top: 446 px; font size: 15 pt) stayed on-screen until participants pressed either the "y" or "n" keys to answer yes or no (Fig. 2a and b), else the screen timed out after one minute.

Researchers often use visual masks (e.g., scrambled patterns) to halt continued visual processing via traces in sensory memory. However, we chose not to do so, as we did not want to inadvertently disrupt processing of the presented scenes in an unequal manner across tasks. For example, a scrambled pattern mask may disrupt processing of configural body features (for the role task) but allow processing color features (for the color task) to go largely unhindered. This would make it difficult to interpret performance differences across tasks, as we do throughout the manuscript. We return to these issues in the General Discussion (for more discussion of these issues and other work using brief displays without visual masks to investigate scene processing, see Breitmeyer, 2007, Sanocki et al., 2023).

2.1.4. Exclusions

Three social interaction categories — scaring, listening to, and pulling — were excluded based on the unspeeded norming study described in Study S1 in the Supplementary Material. Specifically, these categories were excluded because they had low response agreement on at least one of the four tasks (i.e., > 2.5 SDs below each task's mean agreement, where 'agreement' is the proportion of responses in which the majority of participants answered the same way). The remaining 24 categories had high average agreement rates across all tasks (Color: 96%; Role: 89%; Harm: 92%; Moral Wrongness (agent only): 90%). We note that results of this study are qualitatively the same whether or not these three categories were excluded from analyses; in other words, results were statistically significant and in the same direction in both cases.

Participants were excluded for failing the comprehension/attention checks, or if they indicated that they completed a similar task (i.e., they responded "yes" to the post-experiment question: "Have you ever completed a HIT containing a similar scenario, perhaps involving the same sorts of questions?"). We also excluded participants who lost > 15% of their trials due to trial-based exclusion criteria. A trial was excluded from the study if response time was < 150 ms (suggesting that the observer was holding down a key) or > 1 min (suggesting a large lapse in attention).[6] In addition, we recorded presentation durations for each trial using standard JavaScript timing functions and excluded trials that did not meet one of the following timing criteria: (i) the image was presented for < 25 ms or > 40 ms (as measured by the browser), rather than for the intended duration of 33 ms, or (ii) the browser refresh rate was measured at < 30 fps or > 500 fps (suggesting display timing issues). Note that although we excluded data for the above reasons, the results reported below do not depend on these trial or participant exclusions; in other words, all effects remained qualitatively the same (were statistically significant in the same direction) regardless of whether data from these excluded trials and participants were included or not — and the same is true for every experiment reported in this paper.

The sample sizes for each condition after exclusions were the following: Color ($n = 23$), Role ($n = 20$), Harm ($n = 21$), and Moral Wrongness ($n = 41$).

2.2. Results

2.2.1. Response times

Despite the speeded presentation, participants responded quite quickly across the different tasks employed. For Color: $M = 1{,}653$ ms ($SD = 620$ ms); Role: $M = 1{,}673$ ms ($SD = 330$ ms); Harm: 859 ms ($SD = 206$ ms); Moral Wrongness: $M = 1{,}559$ ms ($SD = 532$ ms). We note that response times for the Harm task were about twice as fast as the other tasks. Although this difference may at first glance hint at interesting processing distinctions between Harm and the remaining tasks, the nature of the different prompts for each task suggests caution against making definitive conclusions about them. In particular, the Harm judgment — "Was there harm being inflicted?" — did not require the additional step of considering which individual to respond about, as it did for the other tasks. Remarkably, however, when considering these other tasks (i.e., those that involved a judgment about the person on a particular side, left or right), response times for the morally relevant tasks (Role and Moral Wrongness) appeared to be just as fast as for the non-social baseline Color task.

To formally quantify the differences in response times between task conditions, we conducted two complementary statistical analyses: First, we conducted independent-samples *t*-tests to determine which conditions differed from one another. Below, we report uncorrected *p* values (p_{unc}), and *p* values corrected for the six between-task comparisons using the Bonferroni-Holm method (p_{cor}). To complement these analyses, we also calculated two-sample Bayes Factor *t*-tests between each condition and the next (using the *R* package *BayesFactor* with the function *ttestBF* and the default medium prior of *sqrt*(2)/2). We looked for evidence in favor of the null hypothesis of no difference (i.e., BF_{01}), which would be evinced by a $BF_{01} > 1$.

These statistical analyses confirmed the abovementioned patterns: Color, Role, and Moral Wrongness did not significantly differ from one another (Color vs. Role: $t(34.44) = 0.13$, $p_{unc} = .894$, $p_{cor} > .999$, $d = 0.04$, $BF_{01} = 0.30$; Color vs. Moral Wrongness: $t(31.64) = 0.61$, $p_{unc} = .543$, $p_{cor} > .999$, $d = 0.16$, $BF_{01} = 0.31$; Role vs. Moral Wrongness: $t(55.42) = 1.03$, $p_{unc} = .308$, $p_{cor} = .924$, $d = 0.26$, $BF_{01} = 0.38$), but Harm was significantly different from all other task conditions (Harm vs. Color: $t(27.23) = 5.81$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 1.72$, $BF_{01} = 8.71 \times 10^3$; Harm vs. Role: $t(31.64) = 9.42$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 2.96$, $BF_{01} = 5.48 \times 10^8$; Harm vs. Moral Wrongness: $t(57.11) = 7.41$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 1.74$, $BF_{01} = 4.26 \times 10^4$).

These response time findings suggest that morally relevant judgments after brief exposure to a visual event can be made within the span of just a couple of seconds or less, and at least as quickly as for judgments about basic visual features such as what color shirt a person is wearing. However, regardless of how *fast* people make these judgments, our primary question is how *consistent* these judgments are with the categorizations that observers make with unlimited exposure to such visual scenes. Thus, we leave the interesting question about the specific timing of decisional processes for different moral judgments for future work.

In the remainder of the paper, we focus primarily on measures of successful information extraction, although we note that response time

[6] Although this higher-end cutoff for response times was very liberal (i.e., > 1 min, which excluded just nine trials in total), the results were qualitatively the same across all experiments (i.e., statistically significant in the same direction) even with a much more conservative cutoff (e.g., > 5 sec, which would have excluded 3.1% of trials).

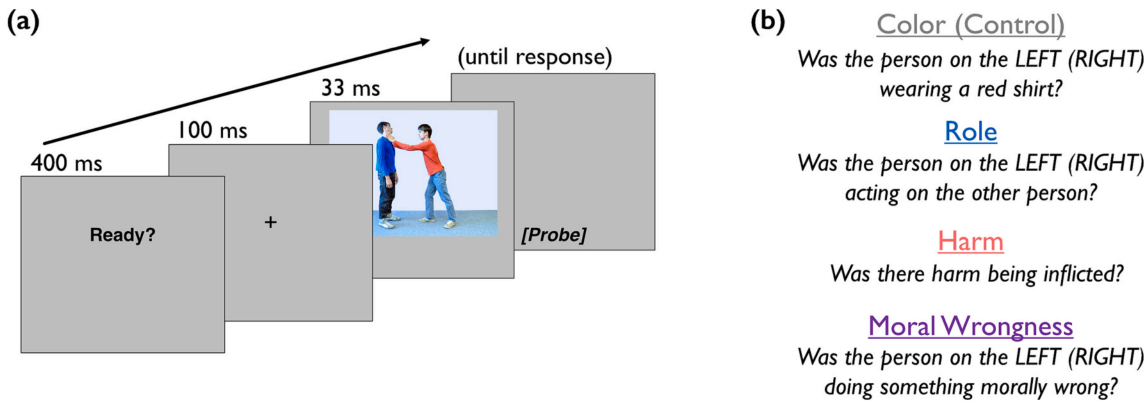

Fig. 2. (a) Participants viewed an image for 33 ms (unmasked) and were then presented with one of four possible probes. (In the norming Study S1, the image and probe appeared at the same time, and both remained on screen until response.) (b) Four probe conditions, shown between-subjects. Participants were asked about the color, role, or moral wrongness of one of the two people in the image (the one on the left or right, asked within-participant), or about whether the interaction involved low or high harm. The same paradigm with different stimuli was used for Studies 3 and 4.

patterns for the different tasks are qualitatively similar across all experiments reported in this paper, and regardless of whether all response times or only ones consistent with unspeeded categorizations are analyzed.

2.2.2. Categorization performance on each task

To measure whether participants extracted morally relevant information about the images at a brief glance in a way that is consistent with unspeeded presentations, we used *d*-prime (d') as our dependent variable of interest. The statistic d' is a bias-free sensitivity measure from signal detection theory capturing, loosely, the difference between signal and noise distributions based on standardized hit and false alarm rates (z-transformed hit rate minus z-transformed false alarm rate; Macmillan & Creelman, 2004). A d' value above zero indicates sensitivity to the information being probed. A response was coded as a 'hit' if a participant responded "yes" in a situation where most unspeeded participants also said "yes" (e.g., for a scene with the red-shirted individual on the right scratching the blue-shirted individual on the left, as in Fig. 2, a 'hit' would be a "yes" response to the question "Was the person on the right wearing a red shirt?"). A response was coded as a 'false alarm' if a participant responded "yes" in a situation where most unspeeded participants said "no" (e.g., for the same red-scratching-blue scene in Fig. 2, a 'false alarm' would be a "yes" response to the question "Was the person on the left acting on the other person?"). Mean hit and false-alarm rates of 0 or 1 were approximated with a standard approach (i.e., zeros were replaced by $1/(2n)$ and ones by $1–1/(2n)$, where n is the number of trials in a group).

We calculated d' for each participant and then tested significance of d' values across participants relative to zero (chance), separately for each task.[7] Remarkably, despite the speeded presentation, participants were able to reliably extract color, role, and harm information, as confirmed by one-sample *t*-tests conducted separately for each task condition (Color: $t(22) = 9.57, p < .001, d = 2.00$; Role: $t(19) = 9.47, p < .001, d = 2.13$; Harm: $t(20) = 8.92, p < .001, d = 1.95$). Furthermore, as hypothesized, participants under brief exposure were able to reliably make moral wrongness judgments ($t(40) = 10.83, p < .001, d = 1.69$), suggesting that they extracted both role and harm information in concert and successfully integrated them to make moral wrongness judgments (Fig. 3).

Post-hoc sensitivity power analyses showed that sample sizes of $n = 41$ (the largest n in any task after exclusions) and $n = 20$ (the smallest n after exclusions) would be sufficient to detect minimum effect sizes of $d = 0.45$ and $d = 0.66$, respectively (one-sample *t*-tests, $\alpha = 0.05$, power $= 0.80$).

2.2.3. Comparison of categorization performance across tasks

To formally quantify differences in sensitivity (d') between task conditions, we conducted the same statistical analyses as reported above for response times: (1) standard independent-samples *t*-tests (with uncorrected and corrected p values), and (2) two-sample Bayes Factor *t*-tests, used to evaluate evidence for the null hypothesis of no difference (which would be evinced by a $BF_{01} > 1$).

Crucially, these tests revealed that Role differed significantly from Moral Wrongness ($t(30.46) = 3.19, p_{unc} = .003, p_{cor} = .013, d = 0.91, BF_{01} = 0.03$), while Harm did not ($t(40.26) = 0.98, p_{unc} = .335, p_{cor} = .335, d = 0.26, BF_{01} = 2.48$) (see also Fig. 3). They also revealed that Color differed significantly from Harm ($t(34.45) = 3.94, p_{unc} < .001, p_{cor} = .002, d = 1.17, BF_{01} = 0.02$) and Moral Wrongness ($t(29.30) = 4.83, p_{unc} < .001, p_{cor} < .001, d = 1.35, BF_{01} < 0.01$), and that the other conditions did not differ significantly from one another: Color vs. Role: $t(39.24) = 2.01, p_{unc} = .052, p_{cor} = .112, d = 0.61, BF_{01} = 0.74$; Role vs. Harm: $t(35.86) = 2.16, p_{unc} = .037, p_{cor} = .112, d = 0.68, BF_{01} = 0.52$.

Post-hoc sensitivity power analyses showed that cross-task comparisons with the two largest sample sizes after exclusions ($n = 41$ and $n = 23$ for Moral Wrongness and Color) and the two smallest sample sizes after exclusions ($n = 20$ and $n = 21$ for Role and Harm) would be sufficient to detect minimum effect sizes of $d = 0.74$ and $d = 0.90$, respectively (two-sample *t*-tests, $\alpha = 0.05$, power $= 0.80$).

2.2.4. Integration of role and harm information for moral wrongness judgments

The fact that Role but not Harm differed from Moral Wrongness suggests that identifying harm served as a perceptual 'bottleneck' (or limiting factor) on performance for the moral wrongness task. This interpretation is reinforced when we look at patterns of responses item-by-item (i.e., social interaction category by social interaction category),

[7] Throughout the paper, we assess the agreement between speeded and unspeeded responses by computing the reliability of d' across participants. This approach offers a measure of performance free from response bias (i.e., the tendency to respond a certain way regardless of stimulus identity), but it necessarily aggregates over stimuli within the same condition for each participant. However, we find comparable results with analyses of individual trial-level 'match' responses (hits and correct rejections) using mixed-effects logistic regression models. While these models do not offer bias-free performance measures, they do offer complementary advantages to our signal-detection-theory approach: they enable generalization of statistical inferences simultaneously across participants and items (social interaction categories) by accounting for both participant- and item-level variability (Barr et al., 2013), and they also deal well with missing trials and unbalanced data. Results of mixed-effects model analyses for Studies 1, 3, and 4 are presented in the Supplementary Material.

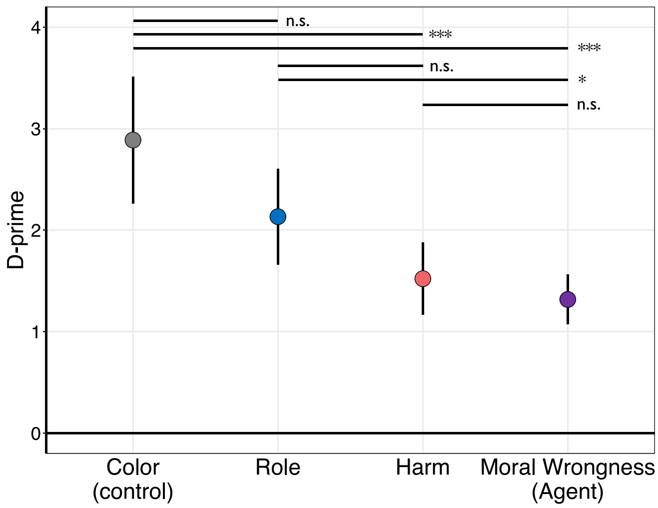

Fig. 3. Participants were above-chance at extracting morally relevant information from visual scenes (where successful and unsuccessful extraction was based on the categorizations that participants with unlimited viewing time made). Points are mean d' values across participants for each judgment condition, and error bars are 95% confidence intervals. Zero is chance performance. Horizontal lines above conditions reflect pairwise significance tests between each condition, corrected for six between-task comparisons using the Bonferroni-Holm method. *** $p < .001$; * $p < .05$; n.s. not significant, $ps > 0.112$.

where we find that responses on the harm task, but not the role task, significantly predict responses on the moral wrongness task (harm: $r(22) = 0.91$, $p < .001$; role: $r(22) = -0.06$, $p = .791$; Fig. 4a). We note that this does not simply mean that harm and moral wrongness are redundant, since making a successful moral wrongness judgment relies on integrating harm information with role information, e.g., even if you know an interaction is harmful, in order to determine whether a given actor is morally wrong you also need to know whether they are the agent or patient of the interaction. It is also worth emphasizing here that the strong relationship between harm and moral wrongness judgments was found in data from two entirely different groups of participants, each focused on only one of the two tasks, further strengthening the evidence for this connection.

Next, we explored whether participants truly integrated role and harm information to make moral wrongness judgments, by conducting a repeated measures ANOVA, with role type (agent vs. patient) and harm level (low- vs. high-harm event) as factors. If role and harm information both informed moral wrongness judgments, we should find an interaction of role type and harm level, such that agents in high-harm interactions are judged as morally wrong more often than agents in low-harm interactions.

Supporting this interpretation, we found significant main effects of role type (agent vs. patient; $F(1,40) = 90.59$, $p < .001$, $\eta_p^2 = 0.87$) and harm level (low- vs. high-harm event; $F(1,40) = 100.2$, $p < .001$, $\eta_p^2 = 0.80$) and a significant interaction between the two ($F(1,40) = 36.75$, $p < .001$, $\eta_p^2 = 0.48$). In post-hoc tests, we found that agents were considered morally wrong significantly more often than patients (68.3% of responses for Agents, 20.2% for Patients, $t(40) = 11.69$, $p < .001$, $d = 1.83$; Fig. 4b), and agents in high-harm interactions were considered morally wrong significantly more often than agents in low-harm interactions (78.2% of responses for Agents in high-harm interactions, 37.6% for Agents in low-harm interactions, $t(40) = 11.58$, $p < .001$, $d = 1.81$; Fig. 4b). Thus, despite the speeded presentation and the fact that participants were never explicitly asked to base their moral judgments on role or harm (as these tasks were between-subject), moral wrongness judgments appeared to leverage information about both.

2.2.5. *Moral wrongness judgments about patients*

Although moral wrongness judgments about agents — the moral transgressors in high-harm events — was the primary focus of our investigation, we also wanted to explore whether moral evaluations about patients would shed light more broadly on the processes involved in moral thin-slicing from the visual world. Not surprisingly, for unspeeded durations (reported in study S1 in the Supplementary Material), we found that patients were considered morally wrong less often than agents for all social interaction categories. Yet we were curious whether this difference required more viewing time to emerge, and more generally, whether there were any differences in judgments of moral wrongness for patients involved in low- versus high-harm events.

To test these questions, we first conducted a mixed-effects ANOVA with harm level (high- vs. low-harm social interactions, within-subject), speed condition (speeded vs. unspeeded, between-subjects), and role (Agent vs. Patient, within-subject) as factors, predicting moral wrongness judgments. The key target of this analysis was the triple interaction of these three factors, probing whether the patterns of moral wrongness judgments for speed condition and harm level varied by role. Indeed, this interaction was significant ($F(1,102) = 21.44$, $p < .001$, $\eta_p^2 = 0.17$; Fig. 4b), prompting us to perform separate mixed-effects ANOVAs on Agent and Patient moral wrongness judgments, with speed condition and harm level as factors.

The Agent moral-wrongness ANOVA showed significant main effects of speed condition ($F(1,102) = 14.05$, $p < .001$, $\eta_p^2 = 0.12$) and harm level ($F(1,102) = 669.60$, $p < .001$, $\eta_p^2 = 0.87$), as well as their interaction ($F(1,102) = 58.74$, $p < .001$, $\eta_p^2 = 0.37$). Notably, under no viewing-time constraints, the proportions of Agent moral wrongness judgments at speed significantly diverged toward floor/ceiling levels, both for low-harm ($t(62.62) = 6.95$, $p < .001$, $d = 1.45$) and high-harm levels ($t(96.01) = 3.18$, $p = .002$, $d = 0.63$) (Fig. 4b). These effects are perhaps not surprising, as the final stimulus set was selected in part based on high agreement for Agent moral wrongness judgments in the unspeeded condition (i.e., Study S1).

By contrast, in the Patient moral-wrongness ANOVA, we found no significant effect of speed condition ($F(1,102) = 3.12$, $p = .080$, $\eta_p^2 = 0.06$), nor an interaction of speed condition and harm level ($F(1,102) = 1.40$, $p = .240$, $\eta_p^2 = 0.01$). Surprisingly, however, we did find a main effect of harm level ($F(1,102) = 40.56$, $p < .001$, $\eta_p^2 = 0.29$). Post-hoc analysis found that patients were attributed moral wrongness significantly more often for participating in a high-harm event (27% of responses) than a low-harm one (8.6% of responses, $t(40) = 6.36$, $p < .001$, $d = 0.62$; Fig. 4b). Indeed, for patient moral wrongness, the highest percentage of responses for a low-harm interaction (*tap*, $M = 13\%$) was still not as high as the lowest percentage for a high-harm interaction (*bite*, $M = 20\%$). In other words, at least in our experimental context, observers judged patients as morally wrong more often when they were involved in high-harm interactions than in low-harm ones — no matter whether they viewed images for a brief glance or had no such constraints on viewing time.

This effect is somewhat reminiscent of so-called 'victim-blaming' effects (Ryan, 1976), although to our knowledge this is the first time that a similar effect has been demonstrated for social interactions that are presented so briefly and that are so generic (involving two similar-looking white men acting in a nondescript setting). However, we note that this effect was not predicted, so it is necessary to interpret it with

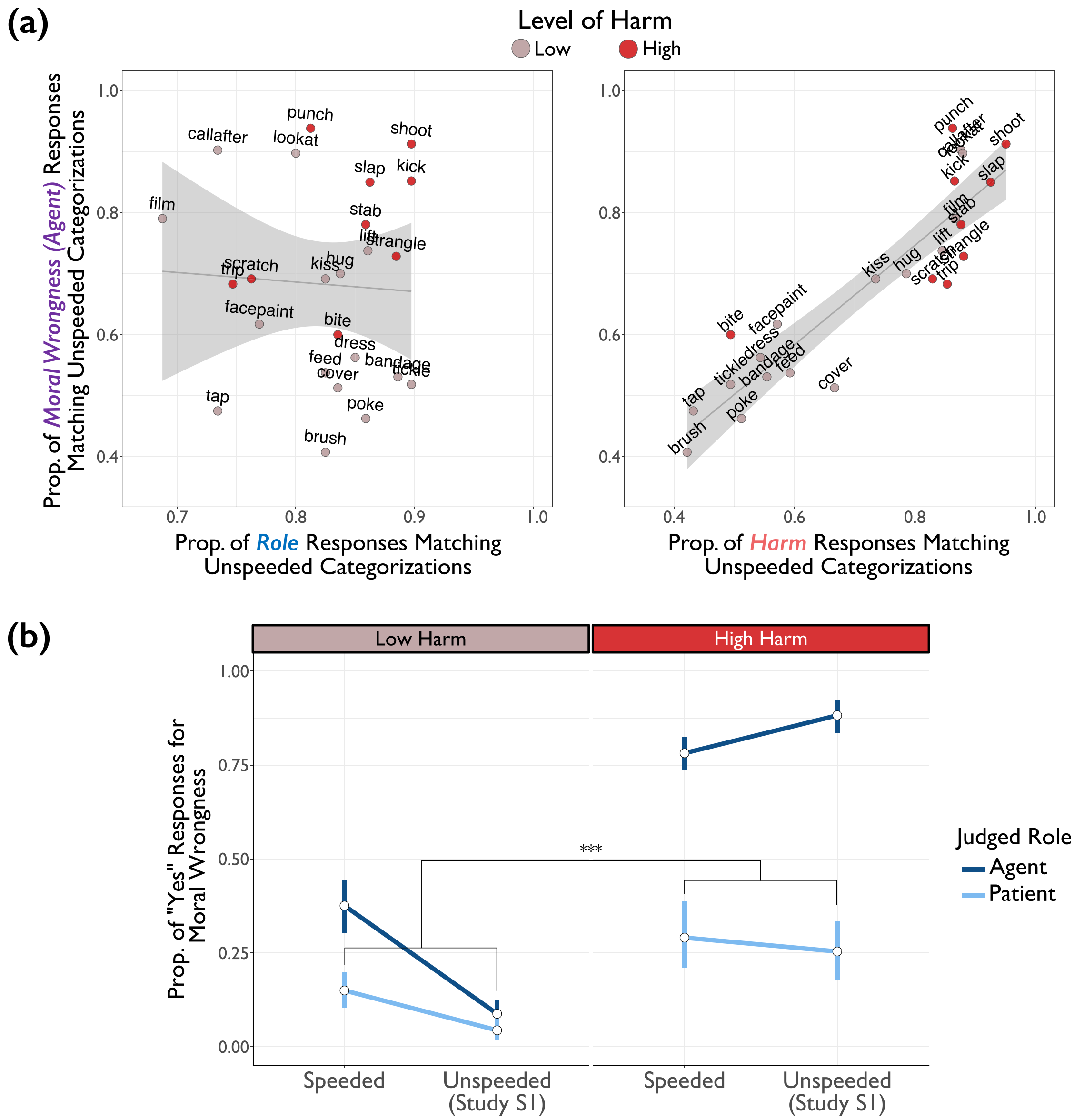

Fig. 4. (a) Pearson correlations across social interactions between the proportion (prop.) of speeded-display Role responses (left panel) or Harm responses (right panel) and the proportion of speeded Moral Wrongness responses (about the Agent) that matched the corresponding unspeeded-display categorizations. There was a significant correlation between Moral Wrongness and Harm performance ($p < .001$) but not Role ($p = .791$). (b) Proportion of "yes" responses for Moral Wrongness, split by Judged Role, Speed, and Harm Level. In speeded displays, high-harm social interactions resulted in more frequent judgments of Agent moral wrongness compared to low-harm interactions, indicating integration of harm and role information. Additionally, regardless of speed, Patients were more often judged as morally wrong in high- vs. low-harm interactions, hinting at a potential 'victim-blaming' bias (although this should be interpreted with caution; see main text for more). Only the statistical significance of the so-called victim-blaming effect is shown in the plot for ease of comparison; for a comprehensive listing of significant effects, see main text. Points are means across participants; error bars are 95% confidence intervals. *** $p < .001$.

caution. We revisit this effect in the General Discussion.[8]

2.2.6. Response biases across tasks

The above analyses highlight the relatively successful extraction of information even under speeded presentation (evinced by higher-than-chance d' values in each task). Even so, it is possible that general tendencies to respond "yes" or "no" would vary by task in ways that could reveal how strategies differ across tasks. To explore this question, we analyzed the 'criterion' or c (-0.5 * (z-transformed hit rate + z-transformed false alarm rate)) separately for each task. The criterion is a decision-making threshold used by an observer to determine whether a stimulus is present or absent (e.g., for the Harm task, whether harm is present or absent; or for the Role task, whether the person on the left was acting on the other person or not). It essentially represents a point on the continuum of evidence where the observer switches from saying "no" to "yes" (or vice versa). A value of zero indicates no bias, negative values indicate a bias to respond "yes", and positive values indicate a bias to respond "no". (Notably, d' — the measure of sensitivity used in our main analyses — is independent from such response biases.) We calculated c for each participant (in each task) and then tested significance of c values across participants relative to zero (chance), separately for each task.

Participants in the Color and Role tasks did not exhibit any significant response biases, as evinced by c values that were not significantly different from zero (Color: $M = 0.08$, 95% CI [−0.04, 0.20], $t(22) = 1.34$, $p = .194$, $d = 0.28$; Role: $M = -0.03$, 95% CI [−0.15, 0.09], $t(19) = 0.54$, $p = .596$, $d = 0.12$). By contrast, participants in the Harm and Moral Wrongness tasks showed a significant bias to respond "yes", as evinced by significant negative c values (Harm: $M = -0.32$, 95% CI [−0.50, −0.14], $t(20) = 3.65$, $p = .002$, $d = 0.80$; Moral Wrongness: $M = -0.24$, 95% CI [−0.42, −0.06], $t(40) = 2.69$, $p = .010$, $d = 0.42$). In other words, participants tended to respond that the actions were harmful and that the agent was doing something morally wrong (even if they were not). This response bias can also be seen visually in Fig. 4a, where the proportion of speeded Harm and Moral Wrongness responses for low-harm social interactions only matched their unspeeded counterparts about half the time (indicating that participants responded "yes" quite often even for these low-harm interactions).

The directionality of these response biases for harm and moral wrongness may be somewhat surprising, given that 15 of the 24 social interaction categories in our study were actually *low*-harm. Nevertheless, despite these response biases, participants still showed significantly positive d' values, demonstrating their capacity to make reliable harm and moral wrongness judgments. Indeed, a major benefit of analyzing the data in this and subsequent studies using signal detection theory (SDT) as opposed to raw accuracies is that in SDT, measurement of sensitivity to information is not affected by the specific criterion an observer sets (hence, why d' is a bias-free measure of sensitivity). For example, two observers may differ in how willing they are to report that an action was harmful (e.g., one may have a c value of -0.25 and so readily report "yes," while the other may have a c value of $+0.25$ and so be more hesitant to do so), even if they are similar in their ability to extract information about harm (e.g., both have d' values of 1.0).

The origin of these response biases is unclear. They may be specific to our study, in the assumptions participants made about what actions would be featured in a psychology experiment. An intriguing possibility is that they also extend to real-world scenarios and observers' assumptions about generic events with similar actors. However, given that sensitivity to morally relevant information was our main goal (and that our measure of such sensitivity, d', is readily interpretable regardless of any response biases), we focus our analyses on d', reporting response-bias analyses for Studies 3 and 4 in the Supplementary Material.

2.3. Discussion

In line with our hypothesis, Study 1 found that the moral judgments about social interactions based on 'thin slices' of viewing time are consistent with the judgments people make under no viewing-time constraints. Observers do this by extracting both role and harm information about the image and combining them to make a moral judgment. We also found that the speed at which they made moral judgments depended on how quickly they extracted harm information.

Finally, we also found an effect somewhat reminiscent of so-called victim-blaming effects (Ryan, 1976), whereby participants judged patients as morally wrong more often when they were involved in high-harm than low-harm interactions, regardless of whether they viewed such scenes only briefly or with unlimited viewing time. This may be the first demonstration that such effects might extend beyond richer social scenarios — e.g., as conveyed through moral vignettes, questionnaire data, thought experiments like the so-called trolley problem (De Freitas et al., 2017; Lerner & Miller, 1978; Niemi & Young, 2016), or unspeeded images or video (Bohner, 2001; Hafer & Begue, 2005) — to generic social interactions. This preliminary finding might suggest that victim blaming could be rooted in a general (albeit slight) bias to assign moral wrongness to patients in high-harm interactions. However, we caution that this effect was unexpected, and other factors such as disparities in social status among participating individuals (e.g., gender, race, economic privilege, power differentials) likely hold a much more significant influence on moral judgments about victims and transgressors. We return to these issues in the General Discussion.

3. Study 2: Temporal evolution of a moral judgment

Just how quick is moral thin-slicing? In other words, what is the minimal amount of time a person needs to view a visual scene to make a well-informed moral judgment, and what determines this speed? Study 2 explored these questions for each of the tasks, by testing performance at 15 different presentation durations ranging from 17 ms to 1500 ms. We planned to test how long an image must be visible before performance at a given speeded presentation is indistinguishable from performance at the unspeeded presentation.

Since harm and role information are the minimal atoms of a moral judgment, at least in our experimental context, the speed of a moral judgment should depend on how quickly an observer can extract both role and harm information. Or put another way, role and harm extraction should serve as the 'temporal bottleneck' on speedy moral judgments, such that a moral judgment cannot be faster than the most slowly extracted information on which the judgment depends.

[8] One might wonder whether such victim-blaming effects were driven by actions involving instruments (e.g., "stabbing"), since for such actions in our stimulus set, both Agent and Patient held identical instruments. Thus, the Patient could appear especially likely to aggress in a dynamic "back-and-forth" with the Agent (e.g., by taking turns stabbing one another). To address this question, we re-ran the Patient moral-wrongness analysis excluding all actions with an instrument, leaving eight low-harm and seven high-harm actions. The results were qualitatively (and even quantitatively) very similar to the analysis with all 24 actions: The Patient moral-wrongness ANOVA again showed no significant effect of speed condition ($F(1,102) = 2.93$, $p = .090$, $\eta_p^2 = 0.06$) nor an interaction of speed condition and harm level ($F(1,102) = 0.01$, $p = .935$, $\eta_p^2 < 0.01$), but it did show a significant main effect of harm level ($F(1,102) = 34.08$, $p < .001$, $\eta_p^2 = 0.25$). Post-hoc analysis found that patients were attributed moral wrongness significantly more often for participating in a high-harm event (25% of responses) than a low-harm one (8.3% of responses, $t(40) = 5.87$, $p < .001$, $d = 0.58$). Thus, although this subset analysis does not on its own fully rule out the possibility that victim-blaming effects might be especially pronounced when the victim appears especially likely to aggress, it does show that the victim-blaming effects we observed are reasonably complex and are not driven by the mere possession of an instrument.

3.1. Method

We collected data from 2,341 participants from Mturk and excluded 350 using the same criteria as in Study 1, leaving 1,991 participants (971 identifying as male, 1,018 as female, 2 not reporting; mean age 36.8, *sd* 11.4, range 18–83, 5 not reporting). We chose the initial sample size to ensure a similar *n* for each task and timing condition compared to Study 1. The design was identical to Study 1, except that the images were presented for one of the following durations, between-subjects: 17, 33, 50, 67, 83, 100, 133, 150, 167, 200, 250, 500, 750, 1000, 1500. See table S1 in Supplementary Material for the exact sample sizes in each condition, after exclusions.

3.2. Results

As in Study 1, we calculated d' values for each participant. We first verified whether we replicated Study 1, by testing for above-chance performance on each task at the 33 ms duration. Again, participants made color, role, harm, and moral wrongness judgments at this duration that were consistent with the judgments that observers made under no viewing-time constraints (Color: $t(26) = 20.24$, $p < .001$, $d = 3.90$; Role: $t(25) = 15.38$, $p < .001$, $d = 3.02$; Harm: $t(27) = 14.45$, $p < .001$, $d = 2.73$; Moral Wrongness: $t(43) = 10.13$, $p < .001$, $d = 1.53$). Post-hoc sensitivity power analyses showed that sample sizes of $n = 44$ (the largest *n* in any task after exclusions) and $n = 26$ (the smallest *n* after exclusions) would be sufficient to detect minimum effect sizes of $d = 0.43$ and $d = 0.57$, respectively (one-sample *t*-tests, $\alpha = 0.05$, power = 0.80).

Furthermore, the general order of task performance was similar to that of Study 1: color performance was greater than that for all other judgments, role was greater than harm and moral wrongness performance, and harm was greater than moral wrongness performance: Color vs. Role: $t(51.00) = 3.71$, $p_{unc} < .001$, $p_{cor} = .002$, $d = 1.02$, $BF_{01} = 0.02$; Color vs. Harm: $t(50.54) = 6.57$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 1.77$, $BF_{01} < 0.01$; Color vs. Moral Wrongness: $t(55.87) = 9.52$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 2.32$, $BF_{01} < 0.01$; Role vs. Harm: $t(49.60) = 2.58$, $p_{unc} = .013$, $p_{cor} = .013$, $d = 0.70$, $BF_{01} = 0.25$; Role vs. Moral Wrongness: $t(55.04) = 5.53$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 1.36$, $BF_{01} < 0.01$; Harm vs. Moral Wrongness: $t(65.63) = 3.25$, $p_{unc} = .002$, $p_{cor} = .004$, $d = 0.77$, $BF_{01} = 0.08$. Post-hoc sensitivity power analyses showed that cross-task comparisons with the two largest sample sizes after exclusions ($n = 44$ and $n = 27$ for Moral Wrongness and Color) and the two smallest sample sizes after exclusions ($n = 26$ and $n = 27$ for Role and Color) would be sufficient to detect minimum effect sizes of $d = 0.70$ and $d = 0.79$, respectively (two-sample *t*-tests, $\alpha = 0.05$, power = 0.80).

Next, to test which speeded durations elicited the same level of performance as the unspeeded presentation, we calculated two-sample Bayes Factor *t*-tests between the unspeeded condition at each speeded duration condition. We looked for evidence in favor of the null hypothesis of no difference (i.e., BF_{01}'s > 1). Evidence for no difference at a given image duration would indicate that the judgment performance at that viewing time reached the level of performance when viewing time was unconstrained. We expected that one of either role or harm extraction should require the same amount of viewing time as moral wrongness, indicating that it served as a perceptual bottleneck on the ability to make the moral wrongness judgment.

Participants detected role more rapidly than they did harm. Specifically, performance on the role task already reached unspeeded levels when the images were presented for just 67 ms, and by 150 ms, performance never dipped below the unspeeded level (Fig. 5; see table S1 in Supplementary Material for statistical tests). In contrast, harm and moral wrongness categorization only reached unspeeded levels when images were presented for about 500 ms. The fact that performance on moral wrongness only reached unspeeded levels when harm did is in line with the perceptual bottleneck account: a moral judgment is only as fast as the slowest extracted piece of information on which it depends. Finally, while Color task performance was always higher than the other tasks at all speeds, it did not reach unspeeded levels until about 1000 ms, which we attribute to the near-ceiling performance of unspeeded participants on the color task. It is possible that since performance on color

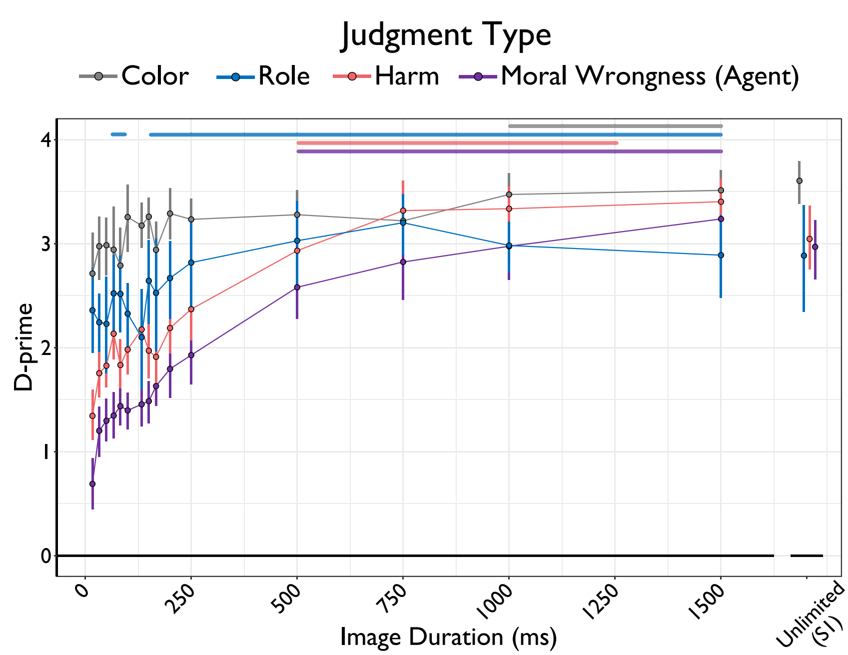

Fig. 5. Images were displayed at durations ranging from 17 ms to 1500 ms. Unspeeded (unlimited viewing time) results from Study S1 are shown for comparison. Points are mean d' values across participants for each judgment condition. Horizontal lines above the plot indicate the durations at which performance on a given task at speeded durations showed no difference from the unspeeded condition (i.e., Bayes Factor > 1). By between 67 and 150 ms, participants performed as well on the Role task as they did under unspeeded presentation, while performance on the Harm and Moral Wrongness tasks did not reach unspeeded levels until at least 500 ms. Error bars are 95% confidence intervals.

was so high for unspeeded presentation, any slight detriment from speeded presentation hurt color performance most.

3.3. Discussion

This study tracked the temporal evolution of a moral judgment, showing that, in line with our predictions, the ability to make well-informed moral wrongness judgments is constrained by the speed at which participants can extract the information on which moral judgment depends. Although participants already extracted role information in a way consistent with unspeeded levels after 67 ms of image presentation, they only did the same for harm information after 500 ms, at which point they could also do the same for moral wrongness. We note that although the order in which each 'atom' of the moral judgment emerged for these images (i.e., role first, then harm) might not generalize to other kinds of images (e.g., ones with more varied scene contexts or viewpoints), this is not as crucial as the fact that (i) different inputs to moral judgment may be extracted at different relative speeds, and (ii) extracting these inputs serves as the temporal bottleneck on making moral judgments under speeded presentation that are consistent with judgments made without viewing-time constraints.

4. Study 3: Causal manipulation of role

If participants are truly using role information to make moral judgments under speeded presentation, then causally manipulating how easy it is to extract role from the images should affect the consistency of moral judgments. To this end, Study 3 showed participants a new set of manipulated images in which the patient of the social interaction leans forward with his limbs outstretched (making him a 'non-prototypical' patient). Typically in social interactions, it is the agent who has such postural characteristics (Hafri et al., 2013), so our goal here was to make it more challenging for an observer to distinguish the agent from the patient than in Studies 1 and 2, in which the actor postures were more prototypical.

4.1. Method

We collected data from 301 participants from Mturk and excluded 37 using the same criteria as Studies 1 and 2, leaving 264 participants (115 identifying as male, 148 as female, 1 not reporting; mean age 37.4, *sd* 11.4, range 18–70). We chose the initial sample size to ensure a similar n for each task compared to the previous studies. The design was identical to Studies 1, except that (i) the images used were different, and (ii) both Speeded (33 ms presentation) and Unspeeded tasks (between-subjects) were included. The primary analyses were performed on responses for the Speeded condition; the Unspeeded condition was only used to evaluate which images should be included in the analyses (see below). The sample sizes for each condition after exclusions were the following: Color (Unspeeded: $n = 33$; Speeded: $n = 29$), Role (Unspeeded: $n = 35$; Speeded: $n = 27$), Harm (Unspeeded: $n = 36$; Speeded: $n = 29$), and Moral Wrongness (Unspeeded: $n = 42$; Speeded: $n = 34$).

The agents in these images were staged identically to those in Studies 1 and 2, but the patient now had similar head orientation, body orientation, extremities, and body lean to the agent, making it more challenging for an observer to distinguish the agent from the patient under speed. Examples can be viewed in Fig. 6a.

Since the aim of this study was to determine whether causally manipulating the difficulty of extracting role information affects judgments of moral wrongness under speed, we needed to ensure that participants could still tell who the agent was when given ample viewing time, i.e., we needed to ensure that our manipulation did not make it seem as though the patient was, in fact, the agent. To this end, we excluded from the analysis any social interactions for which *unspeeded* judgments on non-prototypical patient images (the current study) deviated significantly (> 3.0 *SD*) from the overall distribution of

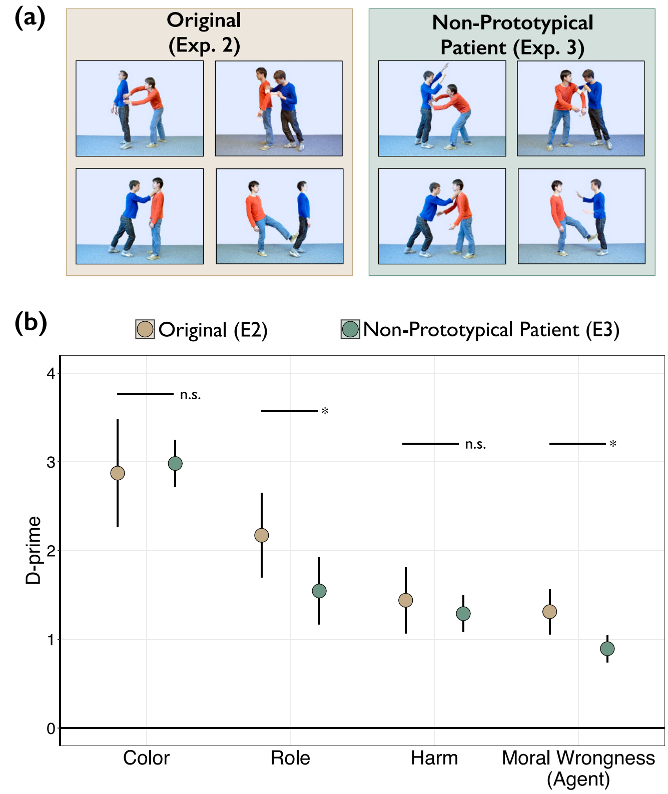

Fig. 6. (a) In Study 3, the Patient's posture was manipulated such that he was a non-prototypical Patient, possessing Agent-like postural features (Hafri et al., 2013). This manipulation was expected to disrupt Role extraction, and to lead to concomitant disruptions to Moral Wrongness extraction. (b) Study 3 results are plotted alongside data from Study 1. Points are mean d' across participants for each judgment condition and error bars are 95% confidence intervals. Horizontal lines above conditions reflect pairwise significance tests between each study. Relative to Study 1, Study 3 showed significant reductions in Role and Moral Wrongness extraction. * $p < .05$; n.s. not significant, $ps > .459$.

categorization differences for each task between the manipulated images in the current study and their non-manipulated counterparts in Study S1 (the original unspeeded norming study, available in the Supplementary Material). This way, the only causally manipulated images that are included are those that are categorized similarly under no viewing-time constraints as the unmanipulated images from Study S1 on all tasks (Color, Role, Harm, and Moral Wrongness); e.g., the red-scratching-blue exemplar would only be included in analyses if the proportion of unspeeded Harm responses for the current study was sufficiently similar to that in Study S1. (However, it is worthwhile to note that these exclusions, if anything, should *hurt* our ability to detect differences in performance based on this role manipulation, as observing such a difference requires that role information which is discernible to a similar degree in the two image categories under ample viewing time is less discernible for the non-prototypical patient images under speeded viewing.)

This exclusion criterion led to the exclusion of two social interactions: *look at* and *tap*. As a sanity check, we confirmed that after this exclusion there were no significant differences in unspeeded performance between Study S1 and the current study on any task (Color: $BF_{01} = 1.78$; Role: $BF_{01} = 3.02$; Harm: $BF_{01} = 2.58$; Moral Wrongness: $BF_{01} = 4.63$). This exclusion procedure left 22 social interactions total in our analysis (13 low-harm and nine high-harm). With that said, we note that none of the results reported here depended on these exclusions, i.e., all effects reported below remain statistically reliable and in the same direction even without excluding these social interactions.

4.2. Results

To test whether manipulating role information causally affected the consistency of moral wrongness judgments, we planned to compare performance on the speeded task on Study 1 (prototypical patient) and the current study (non-prototypical patient). We predicted that with brief viewing time, the role manipulation would impact role performance and thereby moral wrongness performance (since it should depend on role extraction), while the other tasks should remain relatively unaffected. The comparison of Study 3 to Study 1 is justifiable in that the studies did not differ substantially in their basic methods, quality, or variables, aside for the introduction of the image manipulation in Study 3. With that said, we acknowledge the inherently post-hoc nature of such a comparison, which could capture different participant samples, be influenced by differences in timing of the studies, and be open to selective reporting of statistical results. We minimized the chances of this by running the studies on similar Mturk samples at similar times of day, and conducting the most obvious statistical test that would already be expected a priori: comparing the same conditions between studies.

Compared to participants shown the prototypical patient images under speeded displays (Study 1), participants shown the non-prototypical images (the current study) made responses that deviated significantly from unspeeded judgments for the Role task ($t(40.30) = 2.11$, $p = .041$, $d = 0.62$, $BF_{01} = 0.58$) and Moral Wrongness task ($t(65.34) = 2.35$, $p = .022$, $d = 0.53$, $BF_{01} = 0.50$), in line with our hypotheses. At the same time, responses were not statistically different from unspeeded responses for the Color ($t(30.41) = 0.34$, $p = .736$, $d = -0.10$, $BF_{01} = 3.38$) or Harm tasks ($t(33.84) = 0.89$, $p = .377$, $d = 0.26$, $BF_{01} = 2.43$), suggesting that the harm and role tasks relied on non-overlapping visual features (Fig. 6b).

Post-hoc sensitivity power analyses showed that cross-study comparisons with the two largest sample sizes after exclusions ($n = 41$ and $n = 34$ for Moral Wrongness in Studies 1 and 3) and the two smallest sample sizes after exclusions ($n = 20$ and $n = 27$ for Role in Studies 1 and 3) would be sufficient to detect minimum effect sizes of $d = 0.66$ and $d = 0.85$, respectively (two-sample *t*-tests, $\alpha = 0.05$, power $= 0.80$).

4.3. Discussion

In line with our predictions, causally manipulating the accessibility of role information selectively impaired performance on the wrongness task. That is, its effect on moral judgment could be independently isolated from that of harm information on moral judgment. This finding strengthens the earlier conclusion that, ordinarily, role information is independently extracted and then integrated with harm information in order to make speeded moral judgments.

5. Study 4: Causal manipulation of harm

Study 4 aimed to causally manipulate harm information. If participants truly use harm information to make moral judgments under speeded presentation, then, akin to Study 3, causally manipulating how easy it is to extract harm from the images should likewise affect the consistency of moral judgments. To decide on a manipulation, we were informed by the results of Study 2, which suggested that harm information takes longer (about 500 ms) to emerge than other information — presumably because processing fine-grained details in these images takes more time. Thus, we reasoned that making the fine-grained details of the images harder to detect by darkening them would impair harm extraction. Although darkening an image makes all information harder to detect, we predicted that it would lead to less of a decrement in role extraction as compared to harm extraction, since role extraction relies on more global configural features (such as whether individuals are facing toward or away from one another; Hafri et al., 2018, Papeo & Abassi, 2019, Papeo, 2020, Hafri & Firestone, 2021) rather than fine-grained details that are particularly affected by image darkening.[9] To this end, Study 4 showed participants a separate set of manipulated images that were darkened versions of the original images from Study 1.

5.1. Method

We collected data from 292 participants and excluded 27 using the same criteria as in Studies 1–3, leaving 265 participants (160 identifying as male, 105 as female; mean age 34.8, *sd* 12.7, range 18–74, 1 not reporting). We chose the initial sample size to ensure a similar *n* for each task compared to the previous studies. Participants were recruited from the online platform Prolific (for a discussion of this and other online subject pools, see Peer et al., 2017). The design was identical to Study 1, except that the luminance level of the original images was reduced by 80% (see Fig. 7a for examples). This study included both Speeded (33 ms presentation) and Unspeeded tasks (between-subjects), with the analyses of interest being for the Speeded task. Responses for the Unspeeded task were only used to determine what social interaction categories to include (see below). The sample sizes for each condition after exclusions were the following: Color (Unspeeded: $n = 29$; Speeded: $n = 21$), Role (Unspeeded: $n = 30$; Speeded: $n = 23$), Harm (Unspeeded: $n = 30$; Speeded: $n = 25$), and Moral Wrongness (Unspeeded: $n = 62$; Speeded: $n = 45$).

As in Study 3, we excluded from the analysis any social interactions for which *unspeeded* performance on the darkened images (the current study) deviated significantly (> 3.0 *SD*) from the overall distribution of categorization differences between the current study and Study S1 (the unmanipulated images). This led to the exclusion of three social interactions: *strangle, bite,* and *poke*. Since we also planned to compare the current study to Study 3, we additionally excluded the social interactions that were excluded in the previous analyses for Study 3: *look at* and *tap*. After these exclusions, there were no significant differences in unspeeded performance on any condition between the current study and S1 (Color: $BF_{01} = 3.81$; Role: $BF_{01} = 3.33$; Harm: $BF_{01} = 2.87$; Moral Wrongness: $BF_{01} = 4.66$). And even with this larger set of exclusions, there were still no differences in unspeeded performance between Studies 3 and S1 (Color: $BF_{01} = 1.34$; Role: $BF_{01} = 2.90$; Harm: $BF_{01} = 1.89$; Moral Wrongness: $BF_{01} = 4.44$). This exclusion procedure left 19 social interactions total in our analysis (12 low-harm and seven high-harm). However, as in Study 3, we note that none of the results reported here depended on these exclusions, i.e., all effects reported below remain statistically reliable and in the same direction, even without excluding these social interactions.

5.2. Results

We planned to compare performance for the darkened images (the current study) to that for both the original images (Study 1) and non-prototypical patient images (Study 3). For each task, we ran a one-way ANOVA on d' values of the Speeded tasks only, followed by post-hoc two-sample *t*-tests separately comparing data from the current study with Study 1 and with Study 3 (and we corrected for two multiple

[9] One might at first assume that blurring (or low-pass filtering) images would produce the desired effect of making fine-grained details harder to perceive, and thus lead to deficits on harm and moral wrongness judgments without corresponding deficits on role judgments. Indeed, in pilot studies with this manipulation, we observed such effects; however, these effects also carried over to *unspeeded* judgments for harm and moral wrongness, making this potential manipulation problematic. This is likely because the manipulation permanently removes the fine-grained details, making them irrecoverable even without viewing-time constraints. Since our goal was to manipulate the difficulty of extracting harm information *under speed* only, we needed a manipulation that would make harm extraction difficult but recoverable given enough exposure. The darkening manipulation met these criteria and was thus used for the current study.

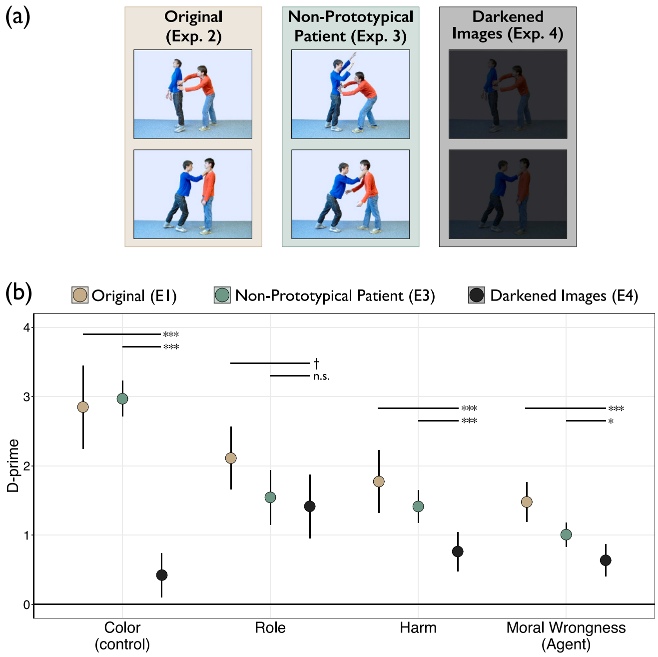

Fig. 7. (a) In Study 4, the luminance level of the original Study 1 images was decreased to a level of 20%. This manipulation was expected to disrupt Harm extraction, and lead to concomitant disruptions to Moral Wrongness extraction. (b) Results from Study 4 are plotted alongside data from Studies 1 and 3, Points are mean d' values across participants for each judgment condition and error bars are 95% confidence intervals. Horizontal lines above conditions reflect pairwise significance tests comparing data from Study 4 to Study 1 and Study 3 (Bonferroni-Holm corrected). Relative to Study 1, Study 4 showed reductions in all judgment conditions. Crucially, although the darkening manipulation of Study 4 reduced role extraction to the levels of Study 3, it had a stronger effect on both Harm and Moral Wrongness categorizations than Study 3; thus, Harm was affected somewhat independently of Role extraction. *** $p < .001$; * $p < .05$; † $p = .060$; n.s. not significant, $p = .663$.

comparisons for each of these two planned cross-study comparisons, using the Bonferroni-Holm method). Once again, we acknowledge the potential limitations of comparing different studies, but took the same steps as in Study 3 to minimize any potential discrepancies between the compared samples, and conducted only the most obvious statistical comparison between the same conditions across studies.

Results can be seen in Fig. 7b. For all tasks, we found either significant or near-significant differences among the three studies: Color ($F(2,70) = 50.00$, $p < .001$, $\eta^2_p = 0.59$), Role ($F(2,67) = 2.82$, $p = .067$, $\eta^2_p = 0.08$), Harm ($F(2,71) = 10.67$, $p < .001$, $\eta^2_p = 0.23$), and Moral Wrongness ($F(2,117) = 13.21$, $p < .001$, $\eta^2_p = 0.18$).

Firstly, post-hoc comparisons revealed that the current study showed a reduction in performance for the baseline Color task relative to the other two studies under speeded presentation (vs. Study 1: $t(33.12) = 7.40$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 2.20$, $BF_{01} < 0.01$; vs. Study 3: $t(42.37) = 12.86$, $p_{unc} < .001$, $p < .001$, $d = 3.69$, $BF_{01} < 0.01$). This result replicates well-established findings of poor color vision under dim illumination (Pokorny et al., 2006). Indeed, analyses of response bias indicate that participants had a generally conservative tendency to respond "no" in this task (i.e., a significantly positive criterion; see Supplementary Material for details). Crucially, this decrement in color extraction performance does not necessarily indicate a problem with the key aim of the darkening manipulation, especially since the moral wrongness task does not ask about color (participants are asked to morally judge the actor on either the left or right, not to judge an actor based on whether they are wearing a red or blue shirt).

Secondly, post-hoc comparisons for role, harm, and moral wrongness revealed that under speeded presentation, the darkening manipulation of this study caused a performance reduction on all three tasks relative to the original images of Study 1: Role ($t(40.92) = 2.24$, $p_{unc} = .030$, $p_{cor} = .060$, $d = 0.68$, $BF_{01} = 0.48$), Harm ($t(32.97) = 3.94$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 1.20$, $BF_{01} = 0.008$), and Moral Wrongness ($t(78.72) = 4.61$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 1.00$, $BF_{01} = 0.001$).

Crucially, and in line with our predictions, although the darkened images of this study reduced Role performance to similar levels as Study 3 (as there was no difference in Role performance between these two studies, $t(45.65) = 0.44$, $p_{unc} = .663$, $p_{cor} = .663$, $d = 0.13$, $BF_{01} = 3.26$), the manipulation had a stronger effect on both Harm and Moral Wrongness judgments relative to Study 3 (Harm: ($t(49.18) = 3.62$, $p_{unc} < .001$, $p_{cor} < .001$, $d = 0.99$, $BF_{01} = 0.02$; Moral Wrongness: $t(75.65) = 2.58$, $p_{unc} = .012$, $p_{cor} = .012$, $d = 0.57$, $BF_{01} = 0.34$). In other words, the darkening manipulation of this study impaired harm and moral wrongness performance more so than the patient role manipulation of Study 3 did. Thus, harm was affected somewhat independently of role extraction, and the greater difficulty in reliably extracting harm information caused by the image manipulation likely led to corresponding difficulties in assigning moral wrongness to the agents in the interaction.

Post-hoc sensitivity power analyses showed that cross-study comparisons with the two largest sample sizes after exclusions ($n = 41$ and $n = 45$ for Moral Wrongness in Studies 1 and 4) and the two smallest sample sizes after exclusions ($n = 20$ and $n = 23$ for Role in Studies 1 and 4) would be sufficient to detect minimum effect sizes of $d = 0.61$ and $d = 0.88$, respectively (two-sample t-tests, $\alpha = 0.05$, power $= 0.80$).

5.3. Discussion

Darkening the original images from Study 1 impaired performance on all tasks. Relative to Study 3, however, the impairment of role detection was comparable even as the impairments of harm and moral wrongness ascriptions were greater, in line with our hypotheses. Thus, making moral judgments based on thin slices may ordinarily rely on extracting harm and role features that are independent, or at least partially so, and then integrating them to make the moral judgment.

6. General discussion

We found that a brief glance at a visual scene is sufficient for observers to extract morally relevant information (event role and harm) and to use this information to make moral judgments that are consistent with those made without viewing-time constraints — evidence for 'moral thin-slicing' from visual observation. Study 1 used a controlled image set and design to show that people are indeed capable of moral thin-slicing. Study 2 presented these same images at various durations and found that people can only make well-informed moral judgments—i.e., ones that are consistent with what they would make with unlimited exposure—once they have extracted role and harm information, which they may do after different durations of viewing time (e.g., in the case of our stimuli, between 67 and 150 ms for role and at least 500 ms for harm). Study 2 suggests that Study 1 provided a temporal snapshot of a process that becomes increasingly accurate the more viewing time is afforded, with performance on the moral wrongness task only approaching unspeeded levels once performance on both the role and harm tasks have also reached this performance plateau. Studies 3 and 4 made visual cues to role and harm harder to detect and found concomitant detriments to the consistency of moral judgments, suggesting that harm and role information affect speeded moral judgments in at least a partially, if not completely, independent manner.

6.1. Moral thin-slicing from visual observation

Our results add to recent perspectives arguing that moral judgments are not always slow and effortful (i.e., a 'System 2' process; Kahneman, 2011), but also are in some cases fast and intuitive — for example, the Theory of Dyadic Morality (Schein & Gray, 2018) or Moral Foundations

Theory (Graham et al., 2013). A uniting factor of such theories is that they rely heavily on empirical data from verbally presented vignettes or scenarios. Our results go beyond this literature by demonstrating that the human visual system in principle can rapidly extract the high-level information on which moral judgment depends, such as role and harm. Furthermore, the visual system not only extracts such information (De Freitas & Alvarez, 2018; Decety & Cacioppo, 2012; Hafri et al., 2013; Hafri et al., 2018; Yoder & Decety, 2014), but it integrates these moral 'atoms' such that they inform moral judgments about events viewed at a brief glance. Notably, this integration was not a given, as there are many cases in other areas of psychology where disparate sources of visual information fail to be integrated toward a common behavioral goal (e.g., for perceiving an object's size, or reorienting in an unfamiliar environment; Rossetti, 1998, Hermer-Vazquez et al., 1999).

Of course, despite the ability to make moral judgments quickly from a brief glance, this does not mean that people do not sometimes slowly deliberate over whether an event was causal, harmful, and so forth, which thought experiments like the trolley problem clearly illustrate (although such scenarios are overly contrived, and deliberately designed to stump readers; De Freitas et al., 2020, De Freitas et al., 2021). Yet the current results suggest that the visual system helps produce a rapid moral judgment when confronted with a range of typical social interactions, circumventing the need to deliberatively mull over this information.

As such, these findings stand in contrast to the characterization of moral judgment as reliant on purely rational inferences about inputs such as causation, harm, etc. without substantive contribution from sensory processing (Martinez & Jaeger, 2016; Olson et al., 2016; Xie et al., 2014). These characterizations suggest that visual processing is involved in moral judgment only in a rudimentary sense, e.g., to recognize objects, their features, and their spatial locations. By contrast, our results add to a growing literature showing that perceptual processing goes beyond such low-level properties, in some cases generating representations of high-level properties such as animacy (Scholl & Gao, 2013), intentionality (Gao et al., 2012), causality (Kominsky & Scholl, 2020; Rolfs et al., 2013), and abstract relations (for a review, see Hafri & Firestone, 2021)— many of which might be readily utilized for moral judgments. Our results also help shed light on how visual processes interface with more abstract modes of cognition: Apparently, information extracted from brief glimpses at visual scenes (in some cases automatically; De Freitas & Alvarez, 2018, Hafri et al., 2018) is sufficient to make reliable moral judgments of the individuals in observed social interactions. It may be that moral thin-slicing occurs because there is a systematic relationship between abstract properties like role and harm and visual stimuli, such that the visual system learns these mappings or has even been naturally selected to do so over evolutionary time, enabling it to extract this information quickly and automatically in the service of other high-level judgments that depend on these inputs, including moral judgment.

6.2. *Relationship to theories of moral judgment*

Our results are related to, yet distinct, from the social intuitionist model of moral judgment (Haidt, 2001) and from more recent pluralistic approaches that rely on intuitions about harm (such as the Theory of Dyadic Morality; Schein & Gray, 2018) or purity (such as Moral Foundations Theory; Graham et al., 2013). At a broad level, our results are in line with the proposal of these theories that moral judgments can be fast, because they often rely on fast and possibly automatic mental processes. Yet whereas these other models find that moral judgments are fast once a basic understanding of the scene has *already* been constructed (e.g., one already knows the event roles and degree of harm), the current work finds that moral judgments are also fast because the mind is fast at understanding a visual scene in the first place, rapidly extracting the atoms of a moral judgment and integrating them to enable a moral judgment.

Of course, this is not to say that moral judgments are not also influenced by information that goes well beyond sensory input, such as contextual information about the social interaction, levels of arousal (Greene et al., 2001), subjective values (Newman et al., 2015), or various heuristics and biases (De Freitas & Johnson, 2018; Haidt et al., 1993). Moreover, our results do not explain why people care about making moral judgment in the first place, nor how they know the moral rules for how to combine pieces of information in order to make well-informed moral judgments (see Curry et al., 2019; De Freitas et al., 2019).

6.3. *Is moral thin-slicing 'accurate'?*

We recognize that broaching the possibility that moral judgments may be 'accurate' is provocative, as such judgments are inherently subjective and involve real social consequences. Keeping these concerns in mind, in one sense, we can say that moral thin-slicing is accurate in that speeded judgments match unspeeded judgments for some morally relevant properties (like event roles) and are a close match for others (harm and moral wrongness). However, just because people show high speeded-unspeeded agreement does not mean that these are good evaluations of the 'true' (but hidden) moral qualities of the individuals depicted in an image. As an example, people also agree on the trustworthiness of faces (Todorov et al., 2009), but this does not necessarily mean they are accurate. Does moral thin-slicing invite similar concerns?

We believe the answer is mixed. On the one hand, visual features indicative of role and harm are likely to be very reliable reflections of the observed social interaction. In other words, if it looks like one agent is harming another, that is likely what is happening. On the other hand, whether someone is truly morally wrong depends on a host of factors other than the immediate social interaction. For instance, the agent may have justifiable reasons for harming the patient, such as self-defense or because they fear that the patient might harm another person, thereby making them less morally wrong than the immediate interaction suggests. Without further contextual information, it is possible that judging an actor's moral wrongness based on an immediate interaction will be globally inaccurate, given all considerations. Another open question for future research is to what extent moral thin-slicing might be influenced by motivated reasoning (Epley & Gilovich, 2016; Kunda, 1990). For instance, might the observers' motivation impact whether they perceive harm being done or not, or how much harm was done, and could this impact how morally wrong the behavior is judged to be? Research could investigate how individual differences in values and motivations interact (or not) with moral thin-slicing in divisive contexts like interpreting potential instances of police brutality or evaluating referee decisions in sporting events. To what extent are observers accurate in these contexts and, if they are inaccurate, does the error originate already in automatic visual processes or is it constrained to more deliberative cognitive processes?

Relatedly, in the current studies we found that, even under speeded presentation, people judged that patients of high-harm interactions were morally wrong more often than patients of low-harm interactions, reminiscent of so-called 'victim blaming' effects (Ryan, 1976). Notably, these effects were reasonably complex, as they were not, for example, solely driven by actions in which both the Agent and Patient possessed an instrument capable of inflicting harm (e.g., a knife in stabbing). It is important to recognize that this effect was unexpected in our studies. Nevertheless, its presence suggests that victim blaming may be a default or universal cognitive bias, albeit a subtle one that may only emerge when other more consequential factors are absent (e.g., differences in gender, race, economic privilege, or power differentials): our stimuli simply involved two White men of equal appearance and age interacting in a neutral scene context. Future work could further explore the factors that contribute to such effects in the context of visual observation.

6.4. Visual stimuli as a tool for evaluating theories of moral judgment

The current work (to our knowledge) is the first to ask how moral impressions are formed from a brief visual glance. To do so, we used a naturalistic but well-controlled set of images — an approach that is necessary for isolating the properties that contribute to moral judgment and how they are extracted. In particular, we precisely manipulated a targeted set of properties, event role and harm level, while simplifying other factors (e.g., by using a neutral scene context).

However, future work could explore whether and how other factors likely to play a role in moral judgment are extracted from visual observation. Crucially, the empirical data from such future investigations could be used to adjudicate between different theories of moral judgment. For example, different theories emphasize certain properties over others (e.g., harm vs. purity; Graham et al., 2013, Schein & Gray, 2018), and some perspectives even recognize distinct types of moral judgment that differentially incorporate aspects of harm, intentionality, norm violations, or general 'badness' evaluations (Malle, 2021). The only one of these factors we manipulated was harm while other factors were kept constant (i.e., the agent was always intentionally acting), which may be one reason why we found that harm and moral wrongness judgments were so highly correlated (Fig. 4a). Future work may jointly manipulate various factors to determine which properties best predict (or which diverge from) speeded moral judgments about visual scenes. It may even be that different properties contribute to moral evaluations at different exposure times (as has been found for inferences about traits from facial appearance; Willis & Todorov, 2006).

More broadly, our study suggests that theories of moral judgment do not have to rely only or even primarily on verbal vignettes; rather, naturalistic visual scenes may be a crucial complementary tool, which has additional advantages in that they may be used across language groups (avoiding biases in the ways that different languages 'package' information about agency and causality; De Freitas et al., 2017) and in populations without full adult proficiency in natural language, such as young children or even non-human primates (Krupenye & Hare, 2018).

6.5. Other limitations and open questions

On a methodological note, while we took various steps to present images for precise, rapid millisecond durations, we did not visually mask images after presentation. This was intentional, since no single mask could have equally disrupted processing of relevant features across all four of our tasks, which differed in various ways (e.g., a simple color feature vs. a configural role feature). Thus, any single mask could have artificially introduced differences in the difficulty of the different tasks. A drawback of this methodological choice is that, although visual processing certainly played a crucial role in enabling speedy moral judgments, we cannot say that our rapid presentations *isolated* feedforward visual processing without also being affected by top-down processing typical of recurrent, attentional, and cognitive processes (Chikkerur et al., 2010; Coltheart, 1980; Milner, 1974). Future work could employ a wide variety of mask types across stimuli and tasks to explore how different types of masks limit processing of different features important for moral judgment (for more discussion of these issues and other work using brief displays without visual masks to investigate scene processing, see Breitmeyer, 2007, Sanocki et al., 2023).

Additionally, while we made our stimuli as controlled as possible in order to isolate the visual processing component of moral judgment, this necessarily simplified the complexities involved when making moral judgments based on observed real-world moral transgressions. For example, in the news and online media, viewing conditions also involve distractions (e.g., music playing in the background), varying informational contexts (e.g., news sources with varying credibility), and pairings between images and text (e.g., headlines, or comments in social media posts). Future work should explore creative ways to test moral thin-slicing in the field that take into account these complexities.

Pairings between image and text are of particular interest, as the linguistic frame may omit, highlight, or minimize certain morally relevant information present in the visual stimulus: for example, changes in the linguistic structure may make the framing more passive ("the customer assaulted the employee" vs. "the employee was assaulted by the customer") or omit certain event components altogether, such as the Agent ("the employee was assaulted"). Such differences in framing can influence how people interpret events conveyed linguistically (De Freitas et al., 2017; Jackendoff, 2010) and how they assign blame for observed events (Fausey & Boroditsky, 2010). Likewise, the way a visual event is recorded and presented visually can alter the linguistic frames used to describe it (e.g., whether the Agent is visible or not; Rissman et al., 2019). Our work lays the foundation for future research on how information from multiple mental systems (vision, cognition, language) are integrated to form moral judgments.

7. Conclusion

Despite the modern rarity with which people are witness to moral transgressions, these transgressions are more accessible than ever thanks to their availability on social media and in the news. Existing work suggests that people make fast moral impressions once they already know what has transpired, e.g., who did what to whom, and whether there was harm (Haidt, 2001). Here, we find that people are also fast at extracting the atoms of moral judgment from a visual scene in the first place and integrating them to decide who is morally wrong, doing so for scenes presented within the blink of an eye (< 100 ms). Our work opens up exciting new avenues for understanding how people form moral judgments from visual observation.

Open practices

We report how we determined sample sizes, all data exclusions, all manipulations, and all measures in the studies. Anonymized trial-level data are publicly available on the Open Science Framework (accessible at https://osf.io/qegjc/). Data were analyzed using R, version 4.0.2 (R Core Team, 2020). Sample sizes were determined before analyses. Study design and analyses were not pre-registered. These statements apply to all studies reported in this manuscript.

CRediT authorship contribution statement

Julian De Freitas: Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Writing – original draft, Writing – review & editing. **Alon Hafri:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Anonymized trial-level data are publicly available on the Open Science Framework and are accessible at https://osf.io/qegjc/.

Acknowledgements

For helpful input, the authors would like to Leyla Isik, Zi Huang, Willis Hole, and Elizabeth Keenan.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2023.104588.

References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4*(7), 267–278.
Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*(2), 256–274.
Ambady, N., & Rosenthal, R. (2006). The 30-sec sale: Using thin-slice judgments to evaluate sales effectiveness. *Journal of Consumer Psychology, 16*(1), 4–13.
Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.
Bohner, G. (2001). Writing about rape: Use of the passive voice and other distancing text features as an expression of perceived responsibility of the victim. *British Journal of Social Psychology, 40*(4), 515–529.
Breitmeyer, B. G. (2007). Visual masking: Past accomplishments, present status, future developments. *Advances in Cognitive Psychology, 3*(1–2), 9–20.
Cameron, C. D., Payne, B. K., Sinnott-Armstrong, W., Scheffer, J. A., & Inzlicht, M. (2017). Implicit moral evaluations: A multinomial modeling approach. *Cognition, 158*, 224–241.
Caruso, E. M., Burns, Z. C., & Converse, B. A. (2016). Slow motion increases perceived intent. *Proceedings of the National Academy of Sciences, 113*(33), 9250–9255.
Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition, 23*(2), 149–178.
Chikkerur, S., Serre, T., Tan, C., & Poggio, T. (2010). What and where: A Bayesian inference theory of attention. *Vision Research, 50*(22), 2233–2247.
Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27*(3), 183–228.
Curhan, J. R., & Pentland, A. (2007). Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first 5 minutes. *Journal of Applied Psychology, 92*(3), 802–811.
Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology, 60*(1), 47–69.
Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108*(2), 353–380.
Cusimano, C. J., Magar, S. T., & Malle, B. F. (2017). Judgment before emotion: People access moral evaluations faster than affective states. In *Proceedings of the 39th annual conference of the cognitive science society, London, UK*.
De Freitas, J., & Alvarez, G. A. (2018). Your visual system provides all the information you need to make moral judgments about generic visual events. *Cognition, 178*, 133–146.
De Freitas, J., & Cikara, M. (2018). Deep down my enemy is good: Thinking about the true self reduces intergroup bias. *Journal of Experimental Social Psychology, 74*, 307–316.
De Freitas, J., & Johnson, S. G. (2018). Optimality bias in moral judgment. *Journal of Experimental Social Psychology, 79*, 149–163.
De Freitas, J., DeScioli, P., Nemirow, J., Massenkoff, M., & Pinker, S. (2017). Kill or die: Moral judgment alters linguistic coding of causality. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 43*(8), 1173–1182.
De Freitas, J., Thomas, K., DeScioli, P., & Pinker, S. (2019). Common knowledge, coordination, and strategic mentalizing in human social life. *Proceedings of the National Academy of Sciences, 116*(28), 13751–13758.
De Freitas, J., Anthony, S. E., Censi, A., & Alvarez, G. (2020). Doubting driverless dilemmas. *Perspectives on Psychological Science, 15*, 1284–1288.
De Freitas, J., Censi, A., Smith, B. W., Di Lillo, L., Anthony, S. E., & Frazzoli, E. (2021). From driverless dilemmas to more practical commonsense tests for automated vehicles. *Proceedings of the National Academy of Sciences, 118*(11), Article e2010202118.
Decety, J., & Cacioppo, S. (2012). The speed of morality: A high-density electrical neuroimaging study. *Journal of Neurophysiology, 108*(11), 3068–3072.
Dennett, D. C. (1989). *The intentional stance*. MIT press.
Epley, N., & Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives, 30*(3), 133–140.
Fausey, C. M., & Boroditsky, L. (2010). Subtle linguistic cues influence perceived blame and financial liability. *Psychonomic Bulletin & Review, 17*(5), 644–650.
Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review, 5*, 5–15.
Gao, T., Scholl, B. J., & McCarthy, G. (2012). Dissociating the detection of intentionality from animacy in the right posterior superior temporal sulcus. *Journal of Neuroscience, 32*(41), 14276–14280.
Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology, Elsevier., 47*, 55–130.
Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry, 23*(2), 206–215.
Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences, 6*(12), 517–523.
Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*(5537), 2105–2108.
Gu, J., Zhong, C.-B., & Page-Gould, E. (2013). Listen to your heart: When false somatic feedback shapes moral behavior. *Journal of Experimental Psychology: General, 142*(2), 307.
Hafer, C. L., & Begue, L. (2005). Experimental research on just-world theory: Problems, developments, and future challenges. *Psychological Bulletin, 131*(1), 128–167.
Hafri, A., Wadhwa, S., & Bonner, M. F. (2022). Perceived distance alters memory for scene boundaries. *Psychological Science, 33*(12), 2040–2058.
Hafri, A., & Firestone, C. (2021). The perception of relations. *Trends in Cognitive Sciences, 25*(6), 475–492.
Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General, 142*(3), 880–905.
Hafri, A., Trueswell, J. C., & Strickland, B. (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition, 175*, 36–52.
Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814–834.
Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Vintage.
Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*(4), 613–628.
Haidt, J., Bjorklund, F., & Murphy, S. (2000). *Moral dumbfounding: When intuition finds no reason* (pp. 191–221). University of Virginia: Unpublished manuscript.
Hall, Z. R., Ahearne, M., & Sujan, H. (2015). The importance of starting right: The influence of accurate intuition on performance in salesperson–customer interactions. *Journal of Marketing, 79*(3), 91–109.
Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature, 370*(6484), 57–59.
Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology, 39*(1), 3–36.
Iliev, R. I., Sachdeva, S., & Medin, D. L. (2012). Moral kinematics: The role of physical factors in moral judgments. *Memory & Cognition, 40*(8), 1387–1401.
Isik, L., Koldewyn, K., Beeler, D., & Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proceedings of the National Academy of Sciences, 114*(43), E9145–E9152.
Jackendoff, R. S. (2010). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press.
Julian, J. B., Keinath, A. T., Muzzio, I. A., & Epstein, R. A. (2015). Place recognition and heading retrieval are mediated by dissociable cognitive systems in mice. *Proceedings of the National Academy of Sciences, 112*(20), 6503–6508.
Kahneman, D. (2011). *Thinking*. Fast and Slow: Macmillan.
Kominsky, J. F., & Scholl, B. J. (2020). Retinotopic adaptation reveals distinct categories of causal perception. *Cognition, 203*, Article 104339.
Krupenye, C., & Hare, B. (2018). Bonobos prefer individuals that hinder others over those that help. *Current Biology, 28*(2), 280–286. e285.
Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480.
Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin, 85*(5), 1030–1051.
Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
Main, K. J., Dahl, D. W., & Darke, P. R. (2007). Deliberative and automatic bases of suspicion: Empirical evidence of the sinister attribution error. *Journal of Consumer Psychology, 17*(1), 59–69.
Malle, B. F. (2021). Moral judgments. *Annual Review of Psychology, 72*, 293–318.
Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry, 25*(2), 147–186.
Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Henry Holt and Co., Inc.
Martinez, L. F., & Jaeger, D. S. (2016). Ethical decision making in counterfeit purchase situations: The influence of moral awareness and moral emotions on moral judgment and purchase intentions. *Journal of Consumer Marketing, 33*(3), 213–223.
Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's mechanical Turk. *Behavior Research Methods, 44*(1), 1–23.
Milner, D., & Goodale, M. (2006). *The visual brain in action*. OUP Oxford.
Milner, P. M. (1974). A model for visual shape recognition. *Psychological Review, 81*(6), 521–535.
Nagel, J., & Waldman, M. (2012). *Force dynamics as a basis for moral intuitions* (Proceedings of the Annual Meeting of the Cognitive Science Society).
Newman, G. E., De Freitas, J., & Knobe, J. (2015). Beliefs about the true self explain asymmetries based on moral judgment. *Cognitive Science, 39*(1), 96–125.
Niemi, L., & Young, L. (2016). When and why we see victims as responsible: The impact of ideology on attitudes toward victims. *Personality and Social Psychology Bulletin, 42*(9), 1227–1242.
Olson, J. G., McFerran, B., Morales, A. C., & Dahl, D. W. (2016). Wealth and welfare: Divergent moral reactions to ethical consumer choices. *Journal of Consumer Research, 42*(6), 879–896.
Papeo, L. (2020). Twos in human visual perception. *Cortex, 132*, 473–478.
Papeo, L., & Abassi, E. (2019). Seeing social events: The visual specialization for dyadic human–human interactions. *Journal of Experimental Psychology: Human Perception and Performance, 45*(7), 877–888.
Patil, I., Calò, M., Fornasier, F., Cushman, F., & Silani, G. (2017). The behavioral and neural basis of empathic blame. *Scientific Reports, 7*(1), 1–14.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153–163.
Peracchio, L. A., & Luna, D. (2006). The role of thin-slice judgments in consumer psychology. *Journal of Consumer Psychology, 16*(1), 25–32.
Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology, 17*(3), 145–171.
Pieters, R., & Wedel, M. (2012). Ad gist: Ad communication in a single eye fixation. *Marketing Science, 31*(1), 59–73.
Pinker, S. (2012). *The better angels of our nature: Why violence has declined*. Penguin Group USA.
Pokorny, J., Lutze, M., Cao, D., & Zele, A. J. (2006). The color of night: Surface color perception under dim illuminations. *Visual Neuroscience, 23*(3–4), 525–530.
R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL https://www.R-project.org/.
Rissman, L., Woodward, A., & Goldin-Meadow, S. (2019). Occluding the face diminishes the conceptual accessibility of an animate agent. *Language, Cognition and Neuroscience, 34*(3), 273–288.
Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology, 23*(3), 250–254.
Rossetti, Y. (1998). Implicit short-lived motor representations of space in brain damaged and healthy subjects. *Consciousness and Cognition, 7*(3), 520–558.
Ryan, W. (1976). *Blaming the victim*. Vintage.
Sanocki, T., Nguyen, T., Shultz, S., & Defant, J. (2023). Novel scene understanding, from gist to elaboration. *Visual Cognition, 31*(3), 188–215.
Schein, C., & Gray, K. (2015). The unifying moral dyad: Liberals and conservatives share the same harm-based moral template. *Personality and Social Psychology Bulletin, 41*(8), 1147–1163.
Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review, 22*(1), 32–70.
Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment. In M. D. Rutherford, & V. A. Kuhlmeier (Eds.), *4629. Social perception: Detection and interpretation of animacy, agency, and intention* (pp. 197–229). MIT Press.
Strohminger, N., & Victor, K. (2018). *The moral psychology of disgust*. Rowman & Littlefield.
Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition, 27*(6), 813–833.
Tsiros, M., Mittal, V., & Ross, W. T., Jr. (2004). The role of attributions in customer satisfaction: A reexamination. *Journal of Consumer Research, 31*(2), 476–483.
Ungerleider, L. G., & Mishkin, M. (1982). In D. J. Ingle, M. A. Goodale, & J. W. Mansfield (Eds.), *Two cortical visual systems. Analysis of visual behavior* (pp. 549–586). Cambridge MA: MIT Press.
Weiner, B. (2000). Attributional thoughts about consumer behavior. *Journal of Consumer Research, 27*(3), 382–387.
Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science, 16*(10), 780–784.
Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*(7), 592–598.
Xie, W., Yu, B., Zhou, X., Sedikides, C., & Vohs, K. D. (2014). Money, moral transgressions, and blame. *Journal of Consumer Psychology, 24*(3), 299–306.
Yoder, K. J., & Decety, J. (2014). Spatiotemporal neural dynamics of moral judgment: A high-density ERP study. *Neuropsychologia, 60*, 39–45.
Young, L., & Saxe, R. (2009). An FMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience, 21*(7), 1396–1405.